

# Approximation of mixed order Sobolev functions on the $d$ -torus – Asymptotics, preasymptotics and $d$ -dependence

Thomas Kühn<sup>a,\*</sup>, Winfried Sickel<sup>b</sup> and Tino Ullrich<sup>c</sup>

<sup>a</sup> Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Germany

<sup>b</sup> Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 2, 07737 Jena, Germany

<sup>c</sup> Hausdorff-Center for Mathematics, Endenicher Allee 62, 53115 Bonn, Germany

*Dedicated to Vladimir N. Temlyakov on the occasion of his 60th birthday*

October 19, 2014

## Abstract

We investigate the approximation of  $d$ -variate periodic functions in Sobolev spaces of dominating mixed (fractional) smoothness  $s > 0$  on the  $d$ -dimensional torus, where the approximation error is measured in the  $L_2$ -norm. In other words, we study the approximation numbers  $a_n$  of the Sobolev embeddings  $H_{\text{mix}}^s(\mathbb{T}^d) \hookrightarrow L_2(\mathbb{T}^d)$ , with particular emphasis on the dependence on the dimension  $d$ . For any fixed smoothness  $s > 0$  we find two-sided estimates for the approximation numbers as a function in  $n$  and  $d$ . We observe super-exponential decay of the constants in  $d$ , if  $n$ , the number of linear samples of  $f$ , is large. In addition, motivated by numerical implementation issues, we also focus on the error decay that can be achieved by approximations using only a few linear samples (small  $n$ ). We present some surprising results for the so-called “preasymptotic” decay and point out connections to the recently introduced notion of quasi-polynomial tractability of approximation problems.

**Keywords** Approximation numbers · Sobolev spaces of mixed smoothness · rate of convergence · preasymptotics ·  $d$ -dependence · quasi-polynomial tractability

**Mathematics Subject Classifications (2000)** 42A10; 41A25; 41A63; 46E35; 65D15

## 1 Introduction

In the present paper we investigate the behavior of the approximation numbers of the embeddings

$$I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d), \quad s > 0, \quad d \in \mathbb{N},$$

where  $H_{\text{mix}}^s(\mathbb{T}^d)$  is the periodic Sobolev space of dominating mixed fractional smoothness  $s$  on the  $d$ -torus  $\mathbb{T}^d$  represented in  $\mathbb{R}^d$  by the cube  $[0, 2\pi]^d$ . The approximation numbers of a

---

\*Corresponding author. Email: kuehn@math.uni-leipzig.de, Tel: +49 341 97 32148. Supported in part by the Spanish Ministerio de Economía y Competitividad (MTM2013-42220-P).

bounded linear operator  $T : X \rightarrow Y$  between two Banach spaces are defined as

$$\begin{aligned} a_n(T : X \rightarrow Y) &:= \inf_{\text{rank } A < n} \sup_{\|x\|_X \leq 1} \|Tx - Ax\|_Y \\ &= \inf_{\text{rank } A < n} \|T - A : X \rightarrow Y\|, \quad n \in \mathbb{N}. \end{aligned} \tag{1.1}$$

They describe the best approximation of  $T$  by finite rank operators. If  $X$  and  $Y$  are Hilbert spaces and  $T$  is compact, then  $a_n(T)$  is the  $n$ th singular number of  $T$ . Moreover, it is well known that for operators in Hilbert spaces the approximation numbers coincide with all other  $s$ -numbers, like Kolmogorov or Gelfand numbers, see e.g. [23, section 11.3].

The first result on the approximation of Sobolev embeddings is due to Kolmogorov [14]. He showed already in 1936 that in the univariate (homogeneous) case with integer smoothness  $m \in \mathbb{N}$  the approximation numbers  $a_n(I : H^m(\mathbb{T}) \rightarrow L_2(\mathbb{T}))$  are given exactly by  $n^{-m}$ . Here we are interested in the multivariate (inhomogeneous) situation, where  $d$  is large, and investigate the approximation numbers  $a_n(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for arbitrary smoothness parameters  $s > 0$ . The spaces  $H_{\text{mix}}^s(\mathbb{T}^d)$  are much smaller than the isotropic spaces  $H^s(\mathbb{T}^d)$ , and often they are considered as a reasonable model for reducing the computational effort in high-dimensional approximation.

In recent years there has been an increasing interest in the approximation of multivariate functions, since many problems, e.g. in finance or quantum chemistry, are modeled in associated function spaces on high-dimensional domains. It has been shown that the functions which have to be approximated often possess a mixed Sobolev regularity, as for instance eigenfunctions of certain Hamilton operators in quantum chemistry, see Yserentant's lecture note [34] and the references given there.

Mityagin [17] was the first who showed the two-sided estimate

$$c_s(d) n^{-s} (\ln n)^{(d-1)s} \leq a_n(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq C_s(d) n^{-s} (\ln n)^{(d-1)s}, \tag{1.2}$$

for  $n \in \mathbb{N}$ ,  $n > 1$ , in a slightly more general setting. Here the constants  $c_s(d)$  and  $C_s(d)$ , depending only on  $d$  and  $s$ , were not explicitly determined. Some more references and comments to the history of (1.2) will be given in Subsection 4.5 below. Our main focus is to clarify, for arbitrary but fixed  $s > 0$ , the dependence of these constants on  $d$ . In fact, it is necessary to fix the norms on the spaces  $H_{\text{mix}}^s(\mathbb{T}^d)$  in advance, since the constants  $c_s(d)$  and  $C_s(d)$  in (1.2) depend on the size of the respective unit balls. Surprisingly, for a collection of quite natural norms (see the next subsection for details) it turns out that we can choose

$$C_s(d) = \left[ \frac{3^d}{(d-1)!} \right]^s.$$

Note that  $C_s(d)$  decays super-exponentially in  $d$ . This observation can be compared to similar results in Bungartz, Griebel [4], Griebel [11], Schwab et al. [26] and Dinh, Ullrich [8], where the authors noticed at least exponential decay of the constants. A more detailed comparison will be made in Subsection 4.5.

Let us ignore the constants  $c_s(d)$  and  $C_s(d)$  for a moment, and fix  $s > 0$ . Then, for arbitrary  $d \in \mathbb{N}$ , the function  $f_d(t) := t^{-s} (\ln t)^{s(d-1)}$  is increasing on  $[1, e^{d-1}]$  and decreasing on  $[e^{d-1}, \infty)$ , hence its maximum on  $[1, \infty)$  is  $\max_{t \geq 1} f_d(t) = f_d(e^{d-1}) = \left(\frac{d-1}{e}\right)^{s(d-1)}$ , which increases super-exponentially in  $d$ . That means, for large  $d$  we have to wait “exponentially long” until the

sequence  $n^{-s}(\ln n)^{(d-1)s}$  decays, and even longer until it becomes less than one. Note that for all norms on  $H_{\text{mix}}^s(\mathbb{T}^d)$  to be considered in this paper, we have  $a_n(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq 1$  for all  $n$ . Consequently, for small values of  $n$  the behavior of  $a_n(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  is not properly reflected by the asymptotic rate  $n^{-s}(\ln n)^{(d-1)s}$ .

This is the reason why we split our investigations into three parts. First, we show that the limit

$$\lim_{n \rightarrow \infty} \frac{n^s \cdot a_n(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{(d-1)s}} = \left[ \frac{2^d}{(d-1)!} \right]^s$$

exists, having the same value for various norms. Secondly, for exponentially large  $n$ , we calculate some admissible constants  $c_s(d)$  and  $C_s(d)$ . Finally, we consider the situation for small  $n$  in some detail. In the range  $1 \leq n \leq 4^d$  we observe

$$2^{-s} \left( \frac{1}{2n} \right)^{\overline{2 + \log_2 \left( \frac{s}{\log_2 n + \frac{1}{2}} \right)}} \leq a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left( \frac{e^2}{n} \right)^{\overline{2 + \log_2 d}},$$

see Theorems 4.9, 4.10 below. Here “#” stands for a specific (but natural) norm in  $H_{\text{mix}}^s(\mathbb{T}^d)$ . For large  $d$  this might be relevant for computational issues, since  $4^d$  pieces of information is already too much for any reasonable algorithm.

The paper is organized as follows. In Section 2 we introduce and investigate the Sobolev spaces of interest. Here we are mainly interested in some assertions on equivalent norms and embeddings. In Subsection 2.2 we add a few remarks to isotropic Sobolev spaces and their relation to Sobolev spaces of dominating mixed smoothness. Subsection 2.3 in this section is devoted to some basics on approximation numbers, in particular, in connection with diagonal operators. In Section 3 we study some combinatorial identities and estimates. Section 4 contains our main results. The final Section 5 transfers our approximation results into the recently introduced notion of quasi-polynomial tractability of the respective approximation problems.

**Notation.** As usual,  $\mathbb{N}$  denotes the natural numbers,  $\mathbb{N}_0$  the non-negative integers,  $\mathbb{Z}$  the integers and  $\mathbb{R}$  the real numbers. By  $\mathbb{T}$  we denote the torus, represented by the interval  $[0, 2\pi]$ , where the end points of the interval are identified. For a real number  $a$  we put  $a_+ := \max\{a, 0\}$  and denote by  $[a]$  its greatest integer part. The letter  $d$  is always reserved for the dimension in  $\mathbb{Z}^d$ ,  $\mathbb{R}^d$ ,  $\mathbb{N}^d$ , and  $\mathbb{T}^d$ . For  $0 < p \leq \infty$  and  $x \in \mathbb{R}^d$  we denote  $|x|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}$  with the usual modification for  $p = \infty$ . If  $\alpha \in \mathbb{N}_0^d$  and  $x \in \mathbb{R}^d$  we use  $x^\alpha := \prod_{i=1}^d x_i^{\alpha_i}$ . The symbol  $\#\Omega$  stands for the cardinality of the set  $\Omega$ . If  $X$  and  $Y$  are two Banach spaces, the norm of an element  $x$  in  $X$  will be denoted by  $\|x\|_X$  and the norm of an operator  $A : X \rightarrow Y$  by  $\|A\|_{X \rightarrow Y}$ . The symbol  $X \hookrightarrow Y$  indicates that there is a continuous embedding from  $X$  into  $Y$ . The equivalence  $a_n \sim b_n$  means that there are constants  $0 < c_1 \leq c_2 < \infty$  such that  $c_1 a_n \leq b_n \leq c_2 a_n$  for all  $n \in \mathbb{N}$ .

## 2 Preliminaries

### 2.1 Sobolev spaces of dominating mixed smoothness on the $d$ -torus

All results in this paper are stated for function spaces on the  $d$ -torus  $\mathbb{T}^d$ , which is represented in the Euclidean space  $\mathbb{R}^d$  by the cube  $\mathbb{T}^d = [0, 2\pi]^d$ , where opposite faces are identified. In particular, for functions  $f$  on  $\mathbb{T}$ , we have  $f(x) = f(y)$  whenever  $x - y = 2\pi k$  for some  $k \in \mathbb{Z}$ . These functions can be viewed as  $2\pi$ -periodic in each component.

The space  $L_2(\mathbb{T}^d)$  consists of all (equivalence classes of) measurable functions  $f$  on  $\mathbb{T}^d$  such that norm

$$\|f\|_{L_2(\mathbb{T}^d)} := \left( \int_{\mathbb{T}^d} |f(x)|^2 dx \right)^{1/2}$$

is finite. All information on a function  $f \in L_2(\mathbb{T}^d)$  is encoded in the sequence  $(c_k(f))_k$  of its Fourier coefficients, given by

$$c_k(f) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{T}^d} f(x) e^{-ik \cdot x} dx, \quad k \in \mathbb{Z}^d.$$

Indeed, we have Parseval's identity

$$\|f\|_{L_2(\mathbb{T}^d)}^2 = \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \quad (2.1)$$

as well as

$$f(x) = \frac{1}{(2\pi)^{d/2}} \sum_{k \in \mathbb{Z}^d} c_k(f) e^{ik \cdot x}$$

with convergence in  $L_2(\mathbb{T}^d)$ .

The mixed Sobolev space  $H_{\text{mix}}^m(\mathbb{T}^d)$  of integer smoothness  $m \in \mathbb{N}$  is the collection of all  $f \in L_2(\mathbb{T}^d)$  such that all distributional derivatives  $D^\alpha f$  of order  $\alpha = (\alpha_1, \dots, \alpha_d)$  with  $\alpha_j \leq m$ ,  $j = 1, \dots, d$ , belong to  $L_2(\mathbb{T}^d)$ . It is usually normed by

$$\|f\|_{H_{\text{mix}}^m(\mathbb{T}^d)} := \left( \sum_{|\alpha|_\infty \leq m} \|D^\alpha f\|_{L_2(\mathbb{T}^d)}^2 \right)^{1/2}. \quad (2.2)$$

One can rewrite this definition in terms of Fourier coefficients. Taking  $c_k(D^\alpha f) = (ik)^\alpha c_k(f)$  into account, Parseval's identity (2.1) implies (using the convention  $0^0 = 1$ )

$$\begin{aligned} \|f\|_{H_{\text{mix}}^m(\mathbb{T}^d)}^2 &= \sum_{|\alpha|_\infty \leq m} \left\| \frac{1}{(2\pi)^{d/2}} \sum_{k \in \mathbb{Z}^d} c_k(f) (ik)^\alpha e^{ik \cdot x} \right\|_{L_2(\mathbb{T}^d)}^2 \\ &= \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \left( \sum_{|\alpha|_\infty \leq m} \prod_{j=1}^d |k_j|^{2\alpha_j} \right) \\ &= \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \left( \prod_{j=1}^d \sum_{\alpha=0}^m |k_j|^{2\alpha} \right) = \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \prod_{j=1}^d v_m(k_j)^2, \end{aligned} \quad (2.3)$$

where

$$\omega_m(\ell)^2 = \sum_{n=0}^m |\ell|^{2n}. \quad (2.4)$$

Due to our convention  $0^0 = 1$  we have  $\omega_m(0) = 1$ , moreover  $\omega_m(\pm 1) = m + 1$ . Defining

$$w_m(k) := \prod_{j=1}^d \omega_m(k_j) \quad \text{for } k = (k_1, \dots, k_d) \in \mathbb{Z}^d, \quad (2.5)$$

we obtain

$$\|f|H_{\text{mix}}^m(\mathbb{T}^d)\| = \left[ \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 w_m(k)^2 \right]^{1/2}.$$

We could also have started with the equivalent norm

$$\|f|H_{\text{mix}}^m(\mathbb{T}^d)\|^* := \left( \sum_{\alpha \in \{0, m\}^d} \|D^\alpha f|L_2(\mathbb{T}^d)\|^2 \right)^{1/2}. \quad (2.6)$$

Similarly as above, a reformulation of (2.6) in terms of Fourier coefficients yields

$$\|f|H_{\text{mix}}^m(\mathbb{T}^d)\|^* = \left[ \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \prod_{j=1}^d (1 + |k_j|^{2m}) \right]^{1/2}. \quad (2.7)$$

Inspired by (2.3) and (2.7) we define Sobolev spaces of dominating mixed smoothness of fractional order  $s > 0$  as follows.

**Definition 2.1.** *Let  $s > 0$ . The periodic Sobolev space  $H_{\text{mix}}^s(\mathbb{T}^d)$  is the collection of all  $f \in L_2(\mathbb{T}^d)$  such that  $\|f|H_{\text{mix}}^s(\mathbb{T}^d)\| < \infty$ . Consider the following equivalent norms.*

(i) *The classical norm  $\|f|H_{\text{mix}}^s(\mathbb{T}^d)\|^+$  is defined as*

$$\|f|H_{\text{mix}}^s(\mathbb{T}^d)\|^+ := \left[ \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \prod_{j=1}^d (1 + |k_j|^2)^s \right]^{1/2}, \quad (2.8)$$

(ii) *the modified classical norm  $\|f|H_{\text{mix}}^s(\mathbb{T}^d)\|^*$  is defined as*

$$\|f|H_{\text{mix}}^s(\mathbb{T}^d)\|^* := \left[ \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \prod_{j=1}^d (1 + |k_j|^{2s}) \right]^{1/2}, \quad (2.9)$$

(iii) *and the norm  $\|f|H_{\text{mix}}^s(\mathbb{T}^d)\|^{\#}$  is a further modification defined by*

$$\|f|H_{\text{mix}}^s(\mathbb{T}^d)\|^{\#} := \left[ \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \prod_{j=1}^d (1 + |k_j|)^{2s} \right]^{1/2}.$$

In the sequel we will often use the notation  $H_{\text{mix}}^{s,+}(\mathbb{T}^d)$ ,  $H_{\text{mix}}^{s,*}(\mathbb{T}^d)$ , and  $H_{\text{mix}}^{s,\#}(\mathbb{T}^d)$  to indicate which of these equivalent norms on  $H_{\text{mix}}^s(\mathbb{T}^d)$  we are considering. For integer smoothness  $s = m \in \mathbb{N}$  all three norms are also equivalent to the norm given in (2.2). Moreover, in some special cases we do not only have equivalence, but even equality of the norms, namely

$$\|\cdot|H_{\text{mix}}^1(\mathbb{T}^d)\| = \|\cdot|H_{\text{mix}}^1(\mathbb{T}^d)\|^* = \|\cdot|H_{\text{mix}}^1(\mathbb{T}^d)\|^+$$

and

$$\|\cdot|H_{\text{mix}}^{1/2}(\mathbb{T}^d)\|^* = \|\cdot|H_{\text{mix}}^{1/2}(\mathbb{T}^d)\|^{\#}.$$

Clearly, the size of the unit balls with respect to equivalent norms can be significantly different. Or, in other words, switching from one to another equivalent norm might produce equivalence constants which badly depend on the dimension  $d$ . Since we are interested in

situations where  $d$  is large or even  $d \rightarrow \infty$ , we have to be very careful with these equivalence constants. Therefore, in this context, norm one embeddings are of particular interest and will be very useful. The embeddings given in the next lemma are due the monotonicity of the norms  $\|\cdot\|_p$ , where  $0 < p < \infty$ , except (v), which is a simple consequence of the fact that the square of an integer is larger than its absolute value.

**Lemma 2.2.** *Let  $s > 0$ . The following embeddings have norm one.*

(i) *If  $s \geq 1$ , then*

$$H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s,*}(\mathbb{T}^d),$$

(ii) *if  $1/2 \leq s \leq 1$ , then*

$$H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s,*}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d),$$

(iii) *if  $s \leq 1/2$ , then*

$$H_{\text{mix}}^{s,*}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d),$$

(iv) *if  $s > t$ , then*

$$H_{\text{mix}}^{s,+}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{t,+}(\mathbb{T}^d) \quad , \quad H_{\text{mix}}^{s,*}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{t,*}(\mathbb{T}^d) \quad , \quad H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{t,\#}(\mathbb{T}^d),$$

(v) *and finally,*

$$H_{\text{mix}}^{s,+}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^{s/2,\#}(\mathbb{T}^d).$$

We also have embeddings where  $H_{\text{mix}}^m(\mathbb{T}^d)$  is involved.

**Lemma 2.3.** *Let  $m \in \mathbb{N}$ . Then for all  $f \in H_{\text{mix}}^m(\mathbb{T}^d)$  the following chain of inequalities holds.*

$$\|f\|_{|H_{\text{mix}}^m(\mathbb{T}^d)|^*} \leq \|f\|_{|H_{\text{mix}}^m(\mathbb{T}^d)|} \leq \|f\|_{|H_{\text{mix}}^m(\mathbb{T}^d)|^+} \leq \left(\frac{2^m}{m+1}\right)^{d/2} \|f\|_{|H_{\text{mix}}^m(\mathbb{T}^d)|} \quad (2.10)$$

**Proof.** The first inequality in (2.10) is obvious. The second one is a consequence of (2.3) and (2.4), together with the fact that  $v_m(\ell)^2 \leq (1 + |\ell|^2)^m$  for all  $\ell \in \mathbb{Z}$  and  $m \in \mathbb{N}$ . For the third inequality, it is enough to notice that

$$\|I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{m,+}(\mathbb{T}^d)\|^2 = \sup_{k \in \mathbb{Z}^d} \prod_{j=1}^d \frac{(1 + |k_j|^2)^m}{v_m(k_j)} = \left(\sup_{k \in \mathbb{N}} \frac{(1 + k^2)^m}{1 + k^2 + \dots + k^{2m}}\right)^d$$

and that the function  $f(x) = \frac{(1+x)^m}{1+x+\dots+x^m}$  is decreasing on  $[1, \infty)$ , hence  $f(x) \geq f(1) = \frac{2^m}{m+1}$ . ■

The most convenient norm for our purposes is  $\|\cdot\|_{|H_{\text{mix}}^s(\mathbb{T}^d)|^\#}$ . In almost all combinatorial estimates given below we use this specific norm. Afterwards, with some additional effort, the results are carried over to the less convenient but more important norms  $\|\cdot\|_{|H_{\text{mix}}^s(\mathbb{T}^d)|^+}$ ,  $\|\cdot\|_{|H_{\text{mix}}^s(\mathbb{T}^d)|^*}$  and  $\|\cdot\|_{|H_{\text{mix}}^m(\mathbb{T}^d)|}$ .

## 2.2 Isotropic Sobolev spaces on the $d$ -torus

Let  $m \in \mathbb{N}$ . Then the isotropic Sobolev space  $H^m(\mathbb{T}^d)$  is the collection of all  $f \in L_2(\mathbb{T}^d)$  such that all distributional derivatives  $D^\alpha f$  up to order  $m$  belong to  $L_2(\mathbb{T}^d)$ , i.e., let

$$\|f\|_{H^m(\mathbb{T}^d)} := \left( \sum_{|\alpha|_1 \leq m} \|D^\alpha f\|_{L_2(\mathbb{T}^d)}^2 \right)^{1/2} < \infty.$$

Fractional versions for  $s > 0$  can be introduced by using Fourier coefficients and the norm

$$\|f\|_{H^s(\mathbb{T}^d)}^+ := \left( \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2 \left(1 + \sum_{j=1}^d |k_j|^2\right)^s \right)^{1/2}.$$

Based on these norms it is easy to compare the isotropic Sobolev spaces with the Sobolev spaces of dominating mixed smoothness.

**Lemma 2.4.** *Let  $s > 0$ . Then we have the chain of continuous norm one embeddings*

$$H^{sd}(\mathbb{T}^d) \hookrightarrow H_{\text{mix}}^s(\mathbb{T}^d) \hookrightarrow H^s(\mathbb{T}^d), \quad (2.11)$$

and this is best possible, i.e., for all  $\varepsilon > 0$ ,

$$H^{sd-\varepsilon}(\mathbb{T}^d) \not\subset H_{\text{mix}}^s(\mathbb{T}^d) \not\subset H^{s+\varepsilon}(\mathbb{T}^d).$$

**Proof.** The proof is elementary, so we will omit the details. However, it is of certain interest to note that the embedding operators in (2.11) are always of norm one, i.e.

$$\|I_d : H^{sd,+}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d)\| = \|I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow H^{s,+}(\mathbb{T}^d)\| = 1$$

for all  $s > 0$ , and

$$\|I_d : H^{md}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^m(\mathbb{T}^d)\| = \|I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow H^m(\mathbb{T}^d)\| = 1$$

for all  $m \in \mathbb{N}$ . ■

The mixed space  $H_{\text{mix}}^s(\mathbb{T}^d)$  is much closer to the space on the left-hand side in (2.11) than to the space on the right-hand side. This is indicated by a short look at the behavior of the approximation numbers. It is known, see, e.g., [33, Chapter 2, Theorems 4.1 and 4.2], that

$$a_s(d) n^{-s/d} \leq a_n(I_d : H^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq A_s(d) n^{-s/d}, \quad n \in \mathbb{N}, \quad (2.12)$$

holds for all  $n$  with constants  $a_s(d)$  and  $A_s(d)$ , only depending on  $d$  and  $s$ , and hence

$$a_n(I_d : H^{sd}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \sim n^{-s}.$$

This coincides up to a logarithmic perturbation with the behavior of  $a_n(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ , see (1.2). Roughly speaking, the mixed Sobolev spaces  $H_{\text{mix}}^s(\mathbb{T}^d)$  are much smaller than their isotropic counterparts  $H^s(\mathbb{T}^d)$ . The behavior of the associated approximation numbers is almost the same as in the one-dimensional isotropic case  $H^s(\mathbb{T})$ . From the very beginning this has been a major motivation to consider spaces of dominating mixed smoothness in approximation theory as well as in the field of information based complexity (IBC). We

refer to Babenko [1], Mityagin [17] and Smolyak [31] for early contributions in the framework of approximation theory (these references are also of relevance with respect to (1.2)). More recent results may be found in Temlyakov's monograph [33]. The role of the spaces  $H_{\text{mix}}^s(\mathbb{T}^d)$  in IBC is summarized in the recent series of books by Novak and Woźniakowski [19, 21, 22]. Observe that in IBC the spaces are sometimes called Korobov spaces, see, e.g., [19, pp. 341].

**Remark 2.5.** In [16] we gave a proof of (2.12) with explicit constants  $a_s(d)$  and  $A_s(d)$  for various equivalent norms.

### 2.3 Approximation numbers

If  $\tau = (\tau_n)_{n=1}^\infty$  is a sequence of real numbers with  $\tau_1 \geq \tau_2 \geq \dots \geq 0$ , we define the diagonal operator  $D_\tau : \ell_2 \rightarrow \ell_2$  by  $D_\tau(\xi) = (\tau_n \xi_n)_{n=1}^\infty$ . Recall the definition of the approximation numbers (1.1) already given in the introduction. The following fact concerning approximation numbers of diagonal operators is well-known, see e.g. Pietsch [23, Theorem 11.3.2.], König [15, Section 1.b], Pinkus [25, Theorem IV.2.2], or Novak and Woźniakowski [19, Corollary 4.12]. Comments on the history may be found in Pietsch [24, 6.2.1.3].

**Lemma 2.6.** *Let  $\tau$  and  $D_\tau$  be as above. Then*

$$a_n(D_\tau : \ell_2 \rightarrow \ell_2) = \tau_n, \quad n \in \mathbb{N}.$$

Here the index set of  $\ell_2$  is  $\mathbb{N}$ . We need a modification for arbitrary countable index sets  $J$ . Then the space  $\ell_2(J)$  is the collection of all  $\xi = (\xi_j)_{j \in J}$  such that the norm

$$\|\xi\|_{\ell_2(J)} := \left( \sum_{j \in J} |\xi_j|^2 \right)^{1/2}$$

is finite. Let  $w = (w_j)_{j \in J}$  with  $w_j > 0$  for all  $j \in J$ , and assume that for every  $\delta > 0$  there are only finitely many  $j \in J$  with  $w_j \geq \delta$ . Then the non-increasing rearrangement  $(\tau_n)_{n \in \mathbb{N}}$  of  $(w_j)_{j \in J}$  exists, and  $\lim_{n \rightarrow \infty} \tau_n = 0$ . Defining  $D_w : \ell_2(J) \rightarrow \ell_2(J)$  by  $D_w(\xi) = (w_j \xi_j)_{j \in J}$  for  $\xi \in \ell_2(J)$ , Lemma 2.6 gives

$$a_n(D_w : \ell_2(J) \rightarrow \ell_2(J)) = \tau_n.$$

The preceding identity is scalable in the following sense.

**Lemma 2.7.** *Let  $J$  be a countable index set, let  $w = (w_j)_{j \in J}$  and  $(\tau_n)_{n \in \mathbb{N}}$  be as above. Then, setting  $w^s = (w_j^s)_{j \in J}$ , one has for any  $s > 0$*

$$a_n(D_{w^s} : \ell_2(J) \rightarrow \ell_2(J)) = a_n(D_w : \ell_2(J) \rightarrow \ell_2(J))^s = \tau_n^s.$$

Now we can reduce our problem on embedding operators in function spaces to the considerably simpler context of diagonal operators in sequence spaces, where the index set is  $J = \mathbb{Z}^d$ . To this end, we consider the operators

$$A_s : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow \ell_2(\mathbb{Z}^d) \quad \text{and} \quad B_s : \ell_2(\mathbb{Z}^d) \rightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d)$$

defined by

$$A_s f = (w_s^+(k) c_k(f))_{k \in \mathbb{Z}^d} \quad \text{and} \quad B_s \xi = (2\pi)^{-d/2} \sum_{k \in \mathbb{Z}^d} \frac{\xi_k}{w_s^+(k)} e^{ik \cdot x},$$



where the weights are  $w_s^+(k) := \prod_{j=1}^d (1 + |k_j|^2)^{s/2}$ . Note the semigroup property of these weights, i.e.,  $w_s^+(k) \cdot w_t^+(k) = w_{s+t}^+(k)$ . Furthermore, we put for  $k \in \mathbb{Z}^d$

$$w(k) := \frac{w_{s_1}^+(k)}{w_{s_0}^+(k)}$$

and make use of the associated diagonal operator  $D_w$ . Then the following commutative diagram illustrates the situation quite well in case  $s_0 > s_1 \geq 0$ :

$$\begin{array}{ccc} H_{\text{mix}}^{s_0,+}(\mathbb{T}^d) & \xrightarrow{I_d} & H_{\text{mix}}^{s_1,+}(\mathbb{T}^d) \\ \downarrow A_{s_0} & & \uparrow B_{s_1} \\ \ell_2(\mathbb{Z}^d) & \xrightarrow{D_w} & \ell_2(\mathbb{Z}^d) \end{array}$$

By the definition of the norm  $\|\cdot\|_{H_{\text{mix}}^s(\mathbb{T}^d)}^+$  it is clear that  $A_s$  and  $B_s$  are isometries and  $B_s = A_s^{-1}$ . For the embedding  $I_d : H_{\text{mix}}^{s_0,+}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s_1,+}(\mathbb{T}^d)$  if  $s_0 > s_1 \geq 0$  we obtain the factorization

$$I_d = B_{s_1} \circ D_w \circ A_{s_0}. \quad (2.13)$$

The multiplicativity of the approximation numbers applied to (2.13) implies

$$a_n(I_d) \leq \|A_{s_0}\| a_n(D_w) \|B_{s_1}\| = a_n(D_w) = \tau_n,$$

where  $(\tau_n)_{n=1}^\infty$  is the non-increasing rearrangement of  $(w(k))_{k \in \mathbb{Z}^d}$ . The reverse inequality can be shown analogously. This gives the important identity

$$a_n(I_d) = a_n(D_w) = \tau_n. \quad (2.14)$$

Of course, (2.14) also holds for  $I_d : H_{\text{mix}}^{s_0,\#}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s_1,\#}(\mathbb{T}^d)$  and for  $I_d : H_{\text{mix}}^{s_0,*}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s_1,*}(\mathbb{T}^d)$  with the obvious adaption of the weights. Due to the semigroup property mentioned above and Lemma 2.7 we have in particular the nice properties

$$\begin{aligned} a_n(I_d : H_{\text{mix}}^{s_0,+}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s_1,+}(\mathbb{T}^d)) &= a_n(I_d : H_{\text{mix}}^{s_0-s_1,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \\ &= a_n(I_d : H_{\text{mix}}^{1,+}(\mathbb{T}^d)_{\text{mix}} \rightarrow L_2(\mathbb{T}^d))^{s_0-s_1} \end{aligned} \quad (2.15)$$

and

$$\begin{aligned} a_n(I_d : H_{\text{mix}}^{s_0,\#}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s_1,\#}(\mathbb{T}^d)) &= a_n(I_d : H_{\text{mix}}^{s_0-s_1,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \\ &= a_n(I_d : H_{\text{mix}}^{1,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))^{s_0-s_1}. \end{aligned}$$

For the norm  $\|\cdot\|^*$  the corresponding weights are

$$w_s^*(k) = \prod_{j=1}^d (1 + |k_j|^{2s})^{1/2}.$$

Note, that they do not satisfy the semigroup property, whence a counterpart of (2.15) does not hold.

### 3 Some combinatorics

In most considerations below, a crucial role will be played by the cardinality  $C(r, d)$  of the set

$$\mathcal{N}(r, d) := \left\{ k \in \mathbb{Z}^d : \prod_{j=1}^d (1 + |k_j|) \leq r \right\}, \quad r \geq 1.$$

**Lemma 3.1.** *For  $r \geq 1$  we have*

$$C(r, d) = 1 + \sum_{\ell=1}^{\min\{d, \log_2 r\}} 2^\ell \binom{d}{\ell} A(r, \ell), \quad (3.1)$$

where  $A(r, \ell) := \#\mathcal{M}(r, \ell)$  with

$$\mathcal{M}(r, \ell) = \left\{ k \in \mathbb{N}^\ell : \prod_{j=1}^{\ell} (1 + k_j) \leq r \right\}.$$

**Proof.** The proof is straightforward. The first summand 1 in (3.1) represents the case  $k_1 = \dots = k_d = 0$ . Next we group together those vectors  $k$  having exactly  $\ell$  non-zero components. This explains why the summation is running from 1 to  $\min\{d, \log_2 r\}$ . Of course, we may concentrate on those  $k \in \mathbb{Z}^d$  with nonnegative components. Since we have  $\ell$  non-zero components, this leads to the factor  $2^\ell$ . Finally, the binomial coefficient  $\binom{d}{\ell}$  is just the number of subsets of  $\{1, \dots, d\}$  of cardinality  $\ell$ . ■

Later on we need estimates of the quantities  $A(r, d)$  for all  $r \in \mathbb{N}$ . Obviously we have  $A(r, d) = 0$  for  $1 \leq r < 2^d$ , and  $A(2^d, d) = 1$ . We intend to relate the number  $A(r, \ell)$  to the  $\ell$ -dimensional Lebesgue measure of the set

$$\mathcal{H}_r^\ell := \left\{ x \in \mathbb{R}^\ell : x_j \geq 1, j = 1, \dots, \ell, \prod_{j=1}^{\ell} x_j \leq r \right\} \subset \mathbb{R}^\ell,$$

which denotes the first orthant part of a so-called hyperbolic cross type set in  $\mathbb{R}^\ell$ . In the sequel we will compute its volume denoted by  $v_\ell(r) := \text{vol}_\ell(\mathcal{H}_r^\ell)$ .

**Lemma 3.2.** *Let  $\ell, r \in \mathbb{N}$  and  $r \geq 2^\ell$ . Then we have*

$$(i) \quad 2^\ell v_\ell(r/2^\ell) \leq A(r, \ell) \leq v_\ell(r) \quad (3.2)$$

and (ii)

$$r \left( \frac{(\ln r)^{\ell-1}}{(\ell-1)!} - \frac{(\ln r)^{\ell-2}}{(\ell-2)!} \right) \leq v_\ell(r) \leq r \frac{(\ln r)^{\ell-1}}{(\ell-1)!}, \quad \ell = 2, 3, \dots$$

Moreover, the upper estimate in (ii) holds as well in case  $\ell = 1$ .

**Proof.** For  $k \in \mathbb{N}^\ell$  put  $Q_k := k + [0, 1]^\ell$ . Then it holds on the one hand

$$\left\{ x \in \mathbb{R}^\ell : x_j \geq 2, \prod_{j=1}^{\ell} x_j \leq r \right\} \subset \bigcup_{k \in \mathcal{M}(r, \ell)} Q_{k+1} \quad (3.3)$$

with  $k + 1 := (k_1 + 1, \dots, k_\ell + 1)$ , and

$$\bigcup_{k \in \mathcal{M}(r, \ell)} Q_k \subset \left\{ x \in \mathbb{R}^\ell : x_j \geq 1, \prod_{j=1}^{\ell} x_j \leq r \right\} \quad (3.4)$$

on the other hand. Taking  $\text{vol}_\ell$  in (3.3) and (3.4) together with a change of variable gives (i).

Let us prove (ii) by induction on  $\ell$ . We first define the function

$$f_\ell(r) := r \frac{(\ln r)^{\ell-1}}{(\ell-1)!}$$

and rewrite (ii) as

$$f_\ell(r) - f_{\ell-1}(r) \leq v_\ell(r) \leq f_\ell(r). \quad (3.5)$$

We consider the upper bound first. One easily verifies the right-hand side in (3.5) in case  $\ell = 1$ . For  $\ell \geq 1$  we use the recurrence relation

$$v_{\ell+1}(r) = \int_1^r v_\ell(r/t) dt,$$

which is a consequence of Fubini's theorem. By a change of variable this can be rewritten as

$$v_{\ell+1}(r) = r \int_1^r v_\ell(s) \frac{ds}{s^2}. \quad (3.6)$$

This implies

$$v_{\ell+1}(r) = r \int_1^r v_\ell(s) \frac{ds}{s^2} \leq r \int_1^r f_\ell(s) \frac{ds}{s^2} = f_{\ell+1}(r).$$

Indeed, the substitution  $u = \ln s$  yields

$$r \int_1^r f_\ell(s) \frac{ds}{s^2} = r \int_1^r \frac{s(\ln s)^{\ell-1}}{(\ell-1)! s^2} ds = \frac{r}{(\ell-1)!} \int_0^{\ln r} u^{\ell-1} du = r \frac{(\ln r)^\ell}{\ell!} = f_{\ell+1}(r). \quad (3.7)$$

For the lower bound we first verify the left-hand side in (3.5) in case  $\ell = 2$  by using  $v_2(r) = r \ln r - r + 1$ . The recurrence relation (3.6) together with the induction hypothesis yields

$$v_{\ell+1}(r) = r \int_1^r v_\ell(s) \frac{ds}{s^2} \geq r \int_1^r \left( f_\ell(s) - f_{\ell-1}(s) \right) \frac{ds}{s^2} = f_{\ell+1}(r) - f_\ell(r),$$

where the last identity is a consequence of (3.7). The proof is complete.  $\blacksquare$

**Remark 3.3.** (i) One of the referees brought to our attention that Lemma 3.2,(i) is equivalent to [5, Lem. 2.1, (2.5)] in the special case  $a = 2$ .

(ii) Lemma 3.2,(ii) above shows

$$\frac{\ln r - \ell + 1}{\ln r} \leq \frac{v_\ell(r)(\ell-1)!}{r(\ln r)^{\ell-1}} \leq 1, \quad (3.8)$$

whereas [5, Cor. 3.3] shows

$$\frac{\ln r}{\ln r + \ell} \leq \frac{v_\ell(r)(\ell-1)!}{r(\ln r)^{\ell-1}} \leq \frac{\ln r}{\ln r + \ell - 1}. \quad (3.9)$$

One observes that the upper bound in (3.9) is always sharper than the upper bound in (3.8), whereas the lower bound in (3.8) is sharper than the lower bound in (3.9) if  $\ln r \geq \ell(\ell-1)$  (i.e. in the asymptotic regime). The lower bound in (3.9) is sharper otherwise (i.e. in the preasymptotic regime). These results are close and complementary.

## 4 Approximation numbers of Sobolev embeddings

In this section we will compute, or at least estimate, the approximation numbers of the embedding  $I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$ . The main aim is to prove (1.2) with explicit constants  $c_s(d)$  and  $C_s(d)$ . First we deal with the norm  $\|\cdot\|_{H_{\text{mix}}^s(\mathbb{T}^d)}^\#$ .

### 4.1 The approximation numbers $a_n(I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ for large $n$

For  $s > 0$  we put

$$w_s^\#(k) := \prod_{j=1}^d (1 + |k_j|)^s, \quad k \in \mathbb{Z}^d.$$

Due to Lemma 2.6 and (2.14) we have

$$a_n(I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = \sigma_n, \quad n \in \mathbb{N},$$

where  $(\sigma_n)_{n \in \mathbb{N}}$  denotes the non-increasing rearrangement of  $(1/w_s^\#(k))_{k \in \mathbb{Z}^d}$ . We have

$$\{\sigma_n : n \in \mathbb{N}\} = \{1/w_s^\#(k) : k \in \mathbb{Z}^d\} = \{r^{-s} : r \in \mathbb{N}_0\},$$

that means  $(\sigma_n)_{n \in \mathbb{N}}$  is a piecewise constant sequence. Recall the notation

$$C(r, d) = \#\left\{k \in \mathbb{Z}^d : \prod_{i=1}^d (1 + |k_i|) \leq r\right\} = \#\left\{k \in \mathbb{Z}^d : w_s^\#(k) \leq r^s\right\}.$$

These observations imply the following result.

**Lemma 4.1.** *Let  $s > 0$  and  $r \in \mathbb{N}$ . Then*

$$a_n(I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = r^{-s}, \quad \text{if } C(r-1, d) < n \leq C(r, d).$$

**Remark 4.2.** Of course, without precise information on the behavior of the quantities  $C(r, d)$ , Lemma 4.1 is not very useful for practical purposes. But it provides, at least in principle, complete knowledge on the sequence of approximation numbers  $a_n(I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ . In particular,

$$\begin{aligned} a_1 = 1 &> a_2 = \dots = a_{2d+1} = 2^{-s} \\ &> a_{2d+2} = \dots = a_{4d+1} = 3^{-s} \\ &> a_{4d+2} = \dots = a_{2d^2+4d+1} = 4^{-s} > \dots \end{aligned}$$

Furthermore, for any  $n \in \mathbb{N}$ , we can easily construct optimal algorithms  $S_n$  of rank less than  $n$ . If  $C(r-1, d) < n \leq C(r, d)$ , we choose

$$S_n f(x) := \frac{1}{(2\pi)^{d/2}} \sum_{k \in \mathcal{N}(r-1, d)} c_k(f) e^{ik \cdot x}.$$

In fact, by this construction we get

$$\sup_{\|f\|_{H_{\text{mix}}^s(\mathbb{T}^d)}^\# \leq 1} \|f - S_n f\|_{L_2(\mathbb{T}^d)} = r^{-s} = a_n(I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)).$$

In a next step we determine the asymptotic behavior of the approximation numbers as  $n \rightarrow \infty$ , including the exact dependence on the smoothness parameter  $s$  and the dimension  $d$ . Note that the *existence* of the limit in the following result is not at all obvious a priori.

**Theorem 4.3.** *Let  $s > 0$  and  $d \in \mathbb{N}$ . Then*

$$\lim_{n \rightarrow \infty} \frac{n^s a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{(d-1)s}} = \left[ \frac{2^d}{(d-1)!} \right]^s.$$

**Proof.** Fix  $d \in \mathbb{N}$ . By Lemma 2.7 it is enough to deal with the case  $s = 1$ . For simplicity of notation we write  $a_n := a_n(I_d : H_{\text{mix}}^{1, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ . We have  $a_n = 1/r$  if  $r \in \mathbb{N}$  and  $C(r-1, d) < n \leq C(r, d)$ , see Lemma 4.1. Clearly  $\lim_{r \rightarrow \infty} C(r, d) = \infty$ , moreover the sequence  $n(\ln n)^{-(d-1)}$  is increasing for  $n > e^{d-1}$ . Hence we obtain for sufficiently large  $r \in \mathbb{N}$  the two-sided inequality

$$\frac{C(r-1, d)}{r(\ln C(r-1, d))^{d-1}} \leq \frac{n a_n}{(\ln n)^{d-1}} \leq \frac{C(r, d)}{r(\ln C(r, d))^{d-1}}. \quad (4.1)$$

By (3.1) and (3.2) we have for  $r \geq 2^d$

$$C(r, d) \leq 1 + \sum_{\ell=1}^d \binom{d}{\ell} 2^\ell v_\ell(r). \quad (4.2)$$

From  $C(r, d) \geq C(r, 1) = 2r - 1 \geq r$  for all  $r \in \mathbb{N}$ , we get  $\ln C(r, d) \geq \ln r$ . Taking Lemma 3.2/(ii) into account, we arrive at

$$\frac{C(r, d)}{r(\ln C(r, d))^{d-1}} \leq \frac{1 + \sum_{\ell=1}^d \binom{d}{\ell} 2^\ell r^{\frac{(\ln r)^{\ell-1}}{(\ell-1)!}}}{r(\ln r)^{d-1}} \xrightarrow{r \rightarrow \infty} \frac{2^d}{(d-1)!},$$

since only the last summand contributes to the limit. Together with (4.1) this gives

$$\limsup_{n \rightarrow \infty} \frac{n a_n}{(\ln n)^{d-1}} \leq \frac{2^d}{(d-1)!}.$$

Now let us pass to the estimate from below. By (3.1) and Lemma 3.2/(i),(ii) we have

$$C(r, d) \geq 2^d r \frac{(\ln r - d \ln 2)^{d-1}}{(d-1)!} \cdot \left( 1 - \frac{d-1}{\ln r - d \ln 2} \right) \quad (4.3)$$

for  $r \geq 2^d$ . Next we need a proper upper estimate for  $\ln C(r, d)$ . In fact, if  $r \geq e^{d-1} \geq e^\ell$  we have

$$\frac{(\ln r)^{\ell-1}}{(\ell-1)!} \leq \frac{(\ln r)^d}{d!}.$$

Hence, using Stirling's formula, we can estimate

$$C(r, d) \leq 1 + \sum_{\ell=1}^d \binom{d}{\ell} 2^\ell r \frac{(\ln r)^d}{d!} = 3^d r \frac{(\ln r)^d}{d!} \leq r \left( \frac{3e \ln r}{d} \right)^d. \quad (4.4)$$

This gives, for  $r \geq e^{d-1}$ ,

$$\ln C(r, d) \leq \ln r + d \ln \ln r + d \ln(3e/d). \quad (4.5)$$

Let now  $C(r, d) < n \leq C(r+1, d)$ . Then we have  $a_n = \frac{1}{r+1}$ , and inserting the above inequalities in (4.3) yields

$$\begin{aligned} \frac{na_n}{(\ln r)^{d-1}} &\geq \frac{C(r, d)}{(r+1)(\ln C(r, d))^{d-1}} \\ &\geq \frac{2^d}{(d-1)!} \cdot \frac{r}{r+1} \cdot \left( \frac{\ln r - d \ln 2}{\ln r + d \ln \ln r + d \ln(3e/d)} \right)^{d-1} \cdot \left( 1 - \frac{d-1}{\ln r - d \ln 2} \right) \\ &\xrightarrow{r \rightarrow \infty} \frac{2^d}{(d-1)!}. \end{aligned}$$

This implies

$$\liminf_{n \rightarrow \infty} \frac{na_n}{(\ln n)^{d-1}} \geq \frac{2^d}{(d-1)!},$$

and the proof is complete. ■

**Remark 4.4.** (i) For  $d = 1$ , the mixed space  $H_{\text{mix}}^{s, \#}(\mathbb{T})$  coincides with the isotropic space  $H^{s, \#}(\mathbb{T})$ , with equality of the corresponding norms. Then

$$\lim_{n \rightarrow \infty} n^s a_n(I_d : H^{s, \#}(\mathbb{T}) \rightarrow L_2(\mathbb{T})) = 2^s$$

follows, a result, already proved in [16].

(ii) As a consequence of Stirling's formula we observe for fixed  $s$  that

$$\left[ \frac{2^d}{(d-1)!} \right]^s \asymp \left( \frac{d}{2\pi} \right)^{s/2} \left( \frac{2e}{d} \right)^{ds},$$

where  $a_d \asymp b_d$  means  $\lim_{d \rightarrow \infty} a_d/b_d = 1$ . This shows a super-exponential decay of the constant.

Being interested in explicit constants  $c_s(d), C_s(d)$  in (1.2), we can learn something from Theorem 4.3. Fix  $d \in \mathbb{N}$  and  $s > 0$ . Then for any given  $\varepsilon > 0$  there exists  $n_0 = n_0(\varepsilon, d, s) \in \mathbb{N}$  such that

$$\left[ \frac{2^d}{(d-1)!} \right]^s - \varepsilon \leq \frac{n^s a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{s(d-1)}} \leq \left[ \frac{2^d}{(d-1)!} \right]^s + \varepsilon \quad \text{for all } n \geq n_0.$$

Equivalently, for any given  $n_1 \in \mathbb{N}$  there is a constant  $\lambda = \lambda(n_1)$ ,  $1 < \lambda < \infty$ , such that

$$\frac{1}{\lambda} \left[ \frac{2^d}{(d-1)!} \right]^s \leq \frac{n^s a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{s(d-1)}} \leq \lambda \left[ \frac{2^d}{(d-1)!} \right]^s \quad \text{for all } n \geq n_1.$$

We now aim at controlling the constant  $\lambda(n_1)$  for certain (large) values of  $n_1$ .

**Theorem 4.5.** *Let  $s > 0$  and  $d \in \mathbb{N}$ .*

(i) *Then*

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left[ \frac{3^d}{(d-1)!} \right]^s \frac{(\ln n)^{(d-1)s}}{n^s} \quad \text{if } n \geq 27^d.$$

(ii) On the other hand,

$$a_n(I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq \left[ \frac{3}{d!} \left( \frac{2}{2 + \ln 12} \right)^d \right]^s \frac{(\ln n)^{(d-1)s}}{n^s} \quad \text{if } n > (12e^2)^d.$$

**Proof.** Again, it is enough to deal with the case  $s = 1$ .

For  $r \in \mathbb{N}$  and  $C(r-1, d) < n \leq C(r, d)$  we have  $a_n = 1/r$ , whence  $a_n \leq 1$  for all  $n$ .

*Step 1.* Proof of (i). First recall that  $C(r, d) \geq C(r, 1) \geq r$  (see the previous proof), and that  $n/(\ln n)^{d-1}$  is increasing for  $n > e^{d-1}$ . Similarly as above in (4.4) we have, for all  $n > e^{d-1}$ ,

$$\frac{n a_n}{(\ln n)^{d-1}} \leq \frac{C(r, d)}{r(\ln C(r, d))^{d-1}} \leq \frac{1 + \sum_{\ell=1}^d 2^\ell \binom{d}{\ell} v_\ell(r)}{r(\ln r)^{d-1}} \leq \frac{1 + \sum_{\ell=1}^d 2^\ell \binom{d}{\ell} \frac{(\ln r)^{\ell-1}}{(\ell-1)!}}{(\ln r)^{d-1}}.$$

Since  $\frac{(\ln r)^{\ell-1}}{(\ell-1)!} \leq \frac{(\ln r)^{d-1}}{(d-1)!}$ , we have

$$\sup_{n \geq C(e^d, d)} \frac{n a_n}{(\ln n)^{d-1}} \leq \frac{1}{(d-1)!} \sum_{\ell=0}^d 2^\ell \binom{d}{\ell} = \frac{3^d}{(d-1)!}.$$

Next we give a precise range for  $n$  in which this estimate holds. To this end, we estimate  $C(r, d)$  with  $r = e^d$  from above. The obvious inequality  $x^k/k! \leq e^x$  applied to  $x = \ln r = d$  gives

$$C(e^d, d) \leq r + \sum_{\ell=1}^d 2^\ell \binom{d}{\ell} r \frac{(\ln r)^{\ell-1}}{(\ell-1)!} \leq 3^d r^2 = (3e^2)^d. \quad (4.6)$$

This shows that (i) holds for all  $n > (3e^2)^d$ .

Finally we show that for  $n \geq 27^d$  the upper bound in (i) is non-trivial (i.e.  $< 1$ ). To see this, we use  $(\ln n)^{d-1}/(d-1)! \leq (\ln n)^d/d!$  for  $n > e^d$ , recall that the function  $f_d(t) = t^{-1}(\ln t)^{d-1}$  is decreasing on  $[e^{d-1}, \infty)$ . Applying Stirling's formula and these monotonicity assertions, estimate (i) yields

$$a_n \leq \frac{(3 \ln n)^d}{d! n} \leq \left( \frac{3ed \ln 27}{27d} \right)^d = \left( \frac{e \ln 3}{3} \right)^d.$$

Since  $(e \ln 3)/3 = 0.99544\dots < 1$ , we see that indeed  $a_n < 1$ .

*Step 2.* Let us turn to the estimate from below. Arguing as in (4.6) we find

$$\ln C(r, d) \leq \ln(3^d r^2), \quad r \geq e^d. \quad (4.7)$$

Next we estimate of  $C(r, d)$  from below. We start with formula (4.3)

$$C(r, d) \geq \frac{2^d r}{(d-1)!} \left( \ln(r/2^d) \right)^{d-1} \left( 1 - \frac{d-1}{\ln(r/2^d)} \right).$$

For  $r \geq r_0 := (2e)^d$  and  $C(r, d) < n \leq C(r+1, d)$ , using again the monotonicity of  $f_d$ , this implies

$$\begin{aligned} \frac{n a_n}{(\ln n)^{d-1}} &\geq \frac{C(r, d)}{(r+1)(\ln C(r, d))^{d-1}} \\ &\geq \frac{2^d}{(d-1)!} \cdot \frac{r}{r+1} \cdot \left( \frac{\ln(r/2^d)}{\ln(3^d r^2)} \right)^{d-1} \cdot \left( 1 - \frac{d-1}{\ln(r/2^d)} \right). \end{aligned} \quad (4.8)$$

Concerning the different factors we have, for all  $r \geq r_0$ ,

$$\begin{aligned} \frac{r}{r+1} &\geq \frac{5^d}{5^d+1} \geq \frac{5}{6}, \quad \text{since } 2e > 5, \\ \frac{\ln(r/2^d)}{\ln(3^d r^2)} &\geq \frac{\ln(e^d)}{\ln(12^d e^{2d})} = \frac{1}{2 + \ln 12} \\ 1 - \frac{d-1}{\ln(r/2^d)} &\geq 1 - \frac{d-1}{d} = \frac{1}{d}. \end{aligned}$$

Hence, taking  $\frac{5}{6}(2 + \ln 12) \geq 3$  into account, we arrive at

$$\sup_{n \geq C(r_0, d)} \frac{n a_n}{(\ln n)^{d-1}} \geq \frac{3}{d!} \left( \frac{2}{2 + \ln 12} \right)^d.$$

Since  $C(r_0, d) \leq 3^d r_0^2 = (12e^2)^d$ , the proof of (ii) is finished.  $\blacksquare$

**Remark 4.6.** (i) We can improve on the bound in (ii), if we choose a larger value of  $r_0$ . But then the range of  $n$  for which (ii) holds becomes smaller. Here is an example. Taking  $r_1 := 48^{d/2} > r_0$ , we get

$$\frac{\ln(r_1/2^d)}{\ln(3^d r_1^2)} = \frac{1}{4}, \quad \frac{r}{r+1} \geq \frac{\sqrt{48}}{\sqrt{48}+1} \geq \frac{41}{47} \quad \text{and} \quad 1 - \frac{d-1}{\ln(r/2^d)} \geq 1 - \frac{1}{\ln \sqrt{12}} = 0.766173... \geq \frac{3}{4}.$$

Since  $\frac{41}{47} \cdot \frac{4}{3} > 1$ , we obtain

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq \left[ \frac{3^d (\ln n)^{d-1}}{2^d (d-1)! n} \right]^s \quad \text{if } n > (48)^{d/2}.$$

(ii) Conversely, one can extend the range of  $n$  in (ii) by making  $r_0$  smaller. However, this strategy is limited by our method. Indeed, if  $r \leq 2^d e^{d-1}$ , then for the last factor in (4.8) we have

$$1 - \frac{d-1}{\ln(r/2^d)} \leq 1 - \frac{d-1}{\ln(e^{d-1})} = 0$$

and our estimate (4.8) becomes useless.

## Some “local” improvements

We do not claim that the estimates obtained in Theorem 4.5 are optimal in  $d$  and  $n$ . They can be improved in various ways. But these improvements take place only locally, i.e., for  $n$  taken from a finite interval.

Let  $d \in \mathbb{N}$ , and let  $(\sigma_n)_{n \in \mathbb{N}}$  be the non-increasing rearrangement of  $\left(1/w_s^\#(k)\right)_{k \in \mathbb{Z}^d}$ . Now we estimate  $\sigma_n$  by a tensor trick. This method is very simple and works for any  $d \in \mathbb{N}$ . The best result that can be obtained in this way differs by a log-factor from the exact asymptotic order of  $\sigma_n$ . However, since the resulting constants are quite explicit, it improves on Theorem 4.5, (i) if  $15^d < n < \exp(\sqrt{d/(2\pi)} \cdot 1.5^d)$ , see Remark 4.8 below.

**Lemma 4.7.** *For every  $d \in \mathbb{N}$ , every  $s > 0$  and all  $n \geq 15^d$  it holds*

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \frac{1}{n^s} \left( \frac{2e \ln n}{d} \right)^{sd}. \quad (4.9)$$



**Proof.** Again we concentrate on  $s = 1$ . For arbitrary  $p > 1$  we have

$$\begin{aligned} n\sigma_n^p &\leq \sum_{j=1}^n \sigma_j^p \leq \sum_{j=1}^{\infty} \sigma_j^p = \sum_{k \in \mathbb{Z}^d} \prod_{\ell=1}^d \frac{1}{(1 + |k_\ell|)^p} = \left( \sum_{m \in \mathbb{Z}} \frac{1}{(1 + |m|)^p} \right)^d = \left( 1 + 2 \sum_{m=2}^{\infty} \frac{1}{m^p} \right)^d \\ &\leq \left( 1 + 2 \int_1^{\infty} \frac{dx}{x^p} \right)^d = \left( 1 + \frac{2}{p-1} \right)^d = \left( \frac{p+1}{p-1} \right)^d, \end{aligned}$$

which implies

$$\sigma_n \leq n^{-1/p} \left( \frac{p+1}{p-1} \right)^{d/p} \leq n^{-1/p} \left( \frac{p+1}{p-1} \right)^d.$$

Now we optimize, for given  $n \in \mathbb{N}$ , over the free parameter  $p > 1$ . Note that the map  $p \mapsto \frac{p+1}{p-1}$  is a bijection from the interval  $(1, \infty)$  onto itself. If  $n > e^{d/2}$ , we have  $(2 \ln n)/d > 1$ , and so we can choose  $p > 1$  such that

$$\frac{p+1}{p-1} = \frac{2 \ln n}{d}. \quad (4.10)$$

It remains to estimate the exponent in  $n^{-1/p}$ . We have

$$-\frac{1}{p} = -1 + \frac{p-1}{p+1} \cdot \frac{p+1}{p} \leq -1 + \frac{d}{\ln n},$$

and hence

$$n^{-1/p} \leq n^{-1} \cdot e^d.$$

This implies the desired estimate

$$\sigma_n \leq \frac{1}{n} \left( \frac{2e \ln n}{d} \right)^d \quad \text{for all } n > e^{d/2}.$$

This bound is non-trivial (i.e.  $< 1$ ) for  $n \geq 15^d$ . ■

**Remark 4.8.** (i) In our Theorem 4.5 we got the upper bound (in slightly rewritten form)

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \frac{1}{n^s \ln^s n} \cdot \left( \sqrt{\frac{d}{2\pi}} \frac{(3e \ln n)^d}{d^d} \right)^s.$$

This bound is larger than the bound obtained in (4.9), if and only if

$$n \leq \exp(\sqrt{d/2\pi} \cdot 1.5^d)$$

which is doubly exponential in  $d$  and therefore already huge for mild dimensions  $d$ . So the tensor trick might after all be quite useful, although it cannot give the exact asymptotic rate.

(ii) The first part of this remark explains that the choice of  $p$  in (4.10) is reasonable, since it almost gives the exact asymptotic rate as  $n \rightarrow \infty$ . However, it is not optimal for all  $n$ . This might be seen as follows. We simply fix  $p$  from the very beginning and follow the above argument. The most simple choice is  $p = 2$ . Then we have the exact value of the sum  $\sum_{j=1}^{\infty} \sigma_j^2$  at hand and the outcome is

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left( \frac{1}{n} \left( \frac{\pi^2}{3} - 1 \right)^d \right)^{s/2}, \quad (4.11)$$

for all  $n \in \mathbb{N}$ . Since  $2 < \frac{\pi^2}{3} - 1 < e$  this estimate is of certain use if  $n \geq e^d$ . Now we compare (4.9) and (4.11). It follows

$$\left(\frac{1}{n} \left(\frac{\pi^2}{3} - 1\right)^d\right)^{s/2} \leq \frac{1}{n^s} \left(\frac{2e \ln n}{d}\right)^{sd}$$

if and only if

$$\frac{n^{1/(2d)}}{\ln n} \leq \frac{2e}{d} \left(\frac{\pi^2}{3} - 1\right)^{-1/2}.$$

A sufficient condition is given by

$$\frac{n^{1/(2d)}}{\ln n} \leq \frac{e}{d}.$$

The function  $f(x) := x^{1/(2d)}/\ln x$  is decreasing on  $[1, e^{2d}]$  and increasing on  $[e^{2d}, \infty)$ , and  $f(e^d) = \sqrt{e}/d < e/d$ . Because of  $f(e^{cd}) = (\sqrt{e})^c/(cd) \leq e/d$  if and only if  $c - 2 \ln c \leq 2$  we conclude that

$$\left(\frac{1}{n} \left(\frac{\pi^2}{3} - 1\right)^d\right)^{s/2} \leq \frac{1}{n^s} \left(\frac{2e \ln n}{d}\right)^{sd} \quad \text{if } e^d \leq n \leq e^{c_0 d},$$

where  $c_0$  is the solution of  $c - 2 \ln c = 2$  ( $5.35 < c_0 < 5.36$ ). Hence, (4.11) is better than (4.9) as long as  $e^d \leq n \leq e^{c_0 d}$ . Different choices of  $p$  (e.g.,  $p = 3/2$ ,  $p = 4$ ) lead to different intervals of optimality, we omit further details.

## 4.2 The approximation numbers $a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ for small $n$

For computational issues, the number  $(3e^2)^d$  in Theorem 4.5 might be much too large. We will now focus on estimates for smaller  $n$  and investigate the so-called preasymptotic behavior. To be more precise, we will deal with estimates of  $a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  from above and below in the range  $1 \leq n \leq (d/2)4^d$ .

**Theorem 4.9.** *Let  $s > 0$  and  $d \in \mathbb{N}$ ,  $d \geq 2$ . For all  $1 \leq n \leq \frac{d}{2}4^d$  it holds*

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left(\frac{e^2}{n}\right)^{\frac{s}{2+\log_2 d}}. \quad (4.12)$$

**Proof.** It is enough to consider the case  $s = 1$ . Let  $1 \leq r \leq 2^d$ . Then for  $C(r-1, d) < n \leq C(r, d)$  we have  $a_n = 1/r$ . Let us estimate  $C(r, d)$  in this case. We shall use  $[x]$  to denote the greatest integer part of the real number  $x$ . Starting from (3.1) and using the obvious estimate  $x^k/k! \leq e^x$  applied to  $x = \ln r = d$ , we obtain

$$\begin{aligned} C(r, d) &= 1 + \sum_{\ell=1}^{[\log_2 r]} 2^\ell \binom{d}{\ell} A(r, \ell) \leq 1 + \sum_{\ell=1}^{[\log_2 r]} 2^\ell \binom{d}{\ell} v_\ell(r) \leq 1 + \sum_{\ell=1}^{[\log_2 r]} 2^\ell \binom{d}{\ell} r \frac{(\ln r)^{\ell-1}}{(\ell-1)!} \\ &\leq r^2 \sum_{\ell=0}^{[\log_2 r]} 2^\ell \binom{d}{\ell} \leq r^2 \sum_{\ell=0}^{[\log_2 r]} 2^\ell \frac{d^\ell}{\ell!} \leq r^2 d^{[\log_2 r]} e^2 \leq e^2 r^{2+\log_2 d}. \end{aligned}$$

This gives  $n \leq C(r, d) \leq e^2 r^{2+\log_2 d}$  which implies  $1/r \leq (e^2/n)^{1/(2+\log_2 d)}$ . Therefore we get for all  $n \leq C(r, d)$  the relation

$$a_n \leq \left(\frac{e^2}{n}\right)^{\frac{1}{2+\log_2 d}}.$$

This estimate holds for all  $n \leq C(2^d, d)$ . We claim that  $C(2^d, d) \geq \frac{d}{2}4^d$ , which implies the assertion. We start with some preparation. Obviously, in case  $\ell \geq 2$ , we have

$$\left\{ k \in \mathbb{N}^\ell : k_2 = \dots = k_\ell = 1, \quad (1 + k_1)2^{\ell-1} \leq r \right\} \subset \mathcal{M}(r, \ell).$$

The set of the left hand side has cardinality  $[r2^{-\ell+1}] - 1$ . By interchanging the roles of  $k_1$  with  $k_j$ ,  $j = 2, 3, \dots, \ell$ , we find  $\ell$  subsets of  $\mathcal{M}(r, \ell)$  having only  $(1, \dots, 1)$  in the intersection. This implies

$$A(r, \ell) \geq \ell [r2^{-\ell+1}] - 2\ell + 1, \quad (4.13)$$

which is also true for  $\ell = 1$ . Using Lemma 3.1 we estimate as follows.

$$\begin{aligned} C(2^d, d) &= 1 + \sum_{\ell=1}^d 2^\ell \binom{d}{\ell} A(2^d, \ell) \geq 1 + \sum_{\ell=1}^d 2^\ell \binom{d}{\ell} (\ell 2^{d-\ell+1} - 2\ell + 1) \\ &= 2^{d+1} \sum_{\ell=1}^d \binom{d}{\ell} \ell - 2 \sum_{\ell=1}^d \binom{d}{\ell} \ell 2^\ell + 3^d \\ &= 3^d + 2^{d+1} \sum_{\ell=1}^d \frac{d!}{(d-\ell)!(\ell-1)!} - 2 \sum_{\ell=1}^d 2^\ell \frac{d!}{(d-\ell)!(\ell-1)!} \\ &= 3^d + 2^{d+1} d \sum_{\ell=1}^d \binom{d-1}{\ell-1} - 2d \sum_{\ell=1}^d 2^\ell \binom{d-1}{\ell-1} \\ &= 3^d + 2^{d+1} d 2^{d-1} - 4d 3^{d-1} \\ &= 3^d + d4^d - 4d3^{d-1}. \end{aligned}$$

Hence, we have  $C(2^d, d) \geq 3^d + d4^d - \frac{4}{3}d3^d$ . Note, that  $C(2^d, d) \geq \frac{d}{2}4^d$  holds if

$$3^d + d4^d - \frac{4}{3}d3^d \geq \frac{d}{2}4^d \quad \iff \quad \frac{d}{2}4^d \geq \left(\frac{4}{3}d - 1\right)3^d. \quad (4.14)$$

Of course, (4.14) is true for all  $d \geq 2$ . The proof is complete.  $\blacksquare$

Let us turn to an estimate from below.

**Theorem 4.10.** *Let  $d \in \mathbb{N}$ ,  $d \geq 2$ , and  $s > 0$ . For  $n \geq 2$  we define*

$$\alpha(n, d) := 2 + \log_2 \left( \frac{d}{\log_2 n} + \frac{1}{2} \right). \quad (4.15)$$

*For all  $2 \leq n \leq \frac{d}{2}4^d$  it holds*

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq 2^{-s} (2n)^{-\frac{s}{\alpha(n, d)}}. \quad (4.16)$$

**Proof.** It suffices to deal with the case  $s = 1$ . Let  $2 \leq r \leq 2^d$  such that  $C(r-1, d) < n \leq C(r, d)$ . Furthermore, let  $m \in \mathbb{N}_0$  be determined from

$$2^m \leq r-1 < 2^{m+1}. \quad (4.17)$$

Then (3.1) and (4.13) imply

$$n > C(r-1, d) = 1 + \sum_{\ell=1}^m 2^\ell \binom{d}{\ell} A(r-1, \ell) \geq \sum_{\ell=1}^m 2^\ell \binom{d}{\ell} (\ell 2^{m-\ell+1} - 2\ell + 1) \geq 2^m \sum_{\ell=1}^m \binom{d}{\ell}.$$

Hence

$$n > 2^m \sum_{\ell=1}^m \binom{m}{\ell} \frac{d(d-1)\cdots(d-\ell+1)}{m(m-1)\cdots(m-\ell+1)} \geq 2^m \sum_{\ell=1}^m \binom{m}{\ell} \left(\frac{d}{m}\right)^\ell.$$

Taking the binomial formula into account, this implies

$$n > 2^m \left\{ \left(1 + \frac{d}{m}\right)^m - 1 \right\} \geq \frac{1}{2} 4^m \left(\frac{d+m}{2m}\right)^m. \quad (4.18)$$

Next we apply  $\log_2$  on both sides and obtain

$$1 + \log_2 n > m \left\{ 2 + \log_2 \left(\frac{d}{2m} + \frac{1}{2}\right) \right\} \geq 2m, \quad (4.19)$$

since  $\frac{d}{2m} + \frac{1}{2} \geq 1$ . Together with (4.19) this yields

$$1 + \log_2 n > m \left\{ 2 + \log_2 \left(\frac{d}{\log_2 n} + \frac{1}{2}\right) \right\} = m \cdot \alpha(n, d).$$

Rewriting this inequality we get

$$2^m < 2^{\frac{1+\log_2 n}{\alpha(n, d)}} = (2n)^{\frac{1}{\alpha(n, d)}}.$$

Taking (4.17) into account, we finally conclude

$$a_n = \frac{1}{r} \geq \frac{1}{2^{m+1}} \geq \frac{1}{2} (2n)^{-\frac{1}{\alpha(n, d)}} \quad (4.20)$$

for all  $n$ ,  $C(1, d) < n \leq C(2^d, d)$ , hence, at least for  $2 \leq n \leq \frac{d}{2} 4^d$  (see (4.14)). If  $s \neq 1$  we obtain (4.16) by raising (4.20) to the power  $s$ . This finishes the proof.  $\blacksquare$

**Remark 4.11.** (i) Of course, there remains a gap between the lower bound in (4.16) and the upper bound in (4.12). For simplicity we comment on this gap for  $s = 1$  only. In fact, we have

$$0 < \frac{1}{\alpha(n, d)} - \frac{1}{2 + \log_2 d} < \frac{\log_2 \log_2 n}{4 + 2 \log_2 d}.$$

The gap is very mildly growing in  $n$  (keeping  $d$  fixed). Therefore, our estimates are losing quality when  $n$  increases. Right now we do not have a conjecture about the correct bounds, most probably both, the lower and the upper estimate, can be improved.

(ii) Note that  $\alpha(n, d)$  is decreasing in  $n$ . Hence, on certain smaller intervals of  $n$ , the dependence on  $n$  in  $\alpha(n, d)$  can be removed by simple monotonicity arguments. For instance, since  $\alpha(4^d, d) = 2$  and  $\alpha(2^{2d/3}, d) = 3$ , we get

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq 2^{-(s+s/2)} n^{-s/2}$$

simultaneously for all  $2 \leq n \leq 4^d$ , and

$$a_n(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq 2^{-(s+s/3)} n^{-s/3}$$

simultaneously for all  $2 \leq n \leq 2^{2d/3}$ .

The approximation rate in these examples is much worse than the asymptotic rate  $n^{-s}$  (ignoring the logarithmic factors). This illustrates well that one has to wait exponentially long until one can “see” the correct asymptotic behavior of the approximation numbers.

### 4.3 The approximation numbers of $H_{\text{mix}}^{s,+}(\mathbb{T}^d)$ , $H_{\text{mix}}^{s,*}(\mathbb{T}^d)$ , and $H_{\text{mix}}^m(\mathbb{T}^d)$ in $L_2(\mathbb{T}^d)$

Now we turn to the investigation of  $a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  and  $a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ .

#### 4.3.1 Preparation

For  $s > 0$  and  $k \in \mathbb{Z}^d$  we put

$$w_s^+(k) := \prod_{j=1}^d (1 + |k_j|^2)^{s/2} \quad \text{and} \quad w_s^*(k) := \prod_{j=1}^d (1 + |k_j|^{2s})^{1/2}, \quad (4.21)$$

see (2.8) and (2.9), respectively. Of interest for us are the non-increasing rearrangements of  $(1/w_s^+(k))_{k \in \mathbb{Z}^d}$  and  $(1/w_s^*(k))_{k \in \mathbb{Z}^d}$ . Let

$$C_s^+(r, d) := \#\{k \in \mathbb{Z}^d : w_{s,+}(k) \leq r\} \quad \text{and} \quad C_s^*(r, d) := \#\{k \in \mathbb{Z}^d : w_{s,*}(k) \leq r\},$$

where  $r \geq 1$ ,  $s > 0$  are real numbers. Let us also define the smaller numbers

$$c_s^+(r, d) := \#\{k \in \mathbb{Z}^d : w_{s,+}(k) < r\} \quad \text{and} \quad c_s^*(r, d) := \#\{k \in \mathbb{Z}^d : w_{s,*}(k) < r\}.$$

In contrast to the weights  $w_s^\#$ , we now have no complete overview over all possible values of  $w_s^+(k)$ ,  $w_s^*(k)$  as  $k$  runs through  $\mathbb{Z}^d$ . Therefore, it is impossible to describe the full sequence of approximation numbers  $a_n$ . However, since  $C_s^+((1+r^2)^{s/2}, d) > c_s^+((1+r^2)^{s/2}, d)$  and  $C_s^*((1+r^{2s})^{1/2}, d) > c_s^*((1+r^{2s})^{1/2}, d)$  if  $r \in \mathbb{N}_0$ , we have at least some partial information about the piecewise constant sequence  $a_n$  of approximation numbers.

**Lemma 4.12.** *Let  $s > 0$  and  $r \in \mathbb{N}_0$ .*

(i) *If  $c_s^+((1+r^2)^{s/2}, d) < n \leq C_s^+((1+r^2)^{s/2}, d)$ , then*

$$a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = (1+r^2)^{-s/2}. \quad (4.22)$$

(ii) *If  $c_s^+((1+r^{2s})^{1/2}, d) < n \leq C_s^+((1+r^{2s})^{1/2}, d)$ , then*

$$a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = (1+r^{2s})^{-1/2}.$$

**Proof.** By Lemma 2.6 and the same principles as used in (2.14), it is enough to note that we have  $w_s^+(k) = (1+r^2)^{s/2}$  and  $w_s^*(k) \geq (1+r^{2s})^{1/2}$  for  $k = (r, 0, \dots, 0) \in \mathbb{Z}^d$ .  $\blacksquare$

#### 4.3.2 Some more combinatorics

Since we have only incomplete information on the set of all values attained by the weights  $w_s^+(k)$  and  $w_s^*(k)$ , it is very difficult to establish similar combinatorial identities and sharp estimates as for the weight  $w_s^\#(k)$ . Therefore we decided for a different strategy. For  $\ell \in \mathbb{Z}$ ,  $0 < \varepsilon \leq 1$  and  $d \in \mathbb{N}$  let

$$y_\ell := \frac{1}{1+|\ell|} \quad , \quad \mathcal{Y}_d(\varepsilon) := \left\{ k \in \mathbb{Z}^d : y_{k_1} \cdots y_{k_d} \geq \varepsilon \right\} \quad , \quad Y_d(\varepsilon) := \#\mathcal{Y}_d(\varepsilon).$$

Because of

$$\left\{ k \in \mathbb{Z}^d : y_{k_1} \cdots y_{k_d} \geq \frac{1}{r} \right\} = \left\{ k \in \mathbb{Z}^d : \prod_{j=1}^d (1+|k_j|) \leq r \right\} = \mathcal{N}(r, d)$$

we have  $Y_d(1/r) = C(r, d)$  for all  $r \in \mathbb{N}_0$ . Using (4.1), Theorem 4.3 and a simple monotonicity argument, this implies

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon \cdot Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} = \lim_{r \rightarrow \infty} \frac{C(r, d)}{r(\ln C(r, d))^{d-1}} = \frac{2^d}{(d-1)!}. \quad (4.23)$$

**Lemma 4.13.** *For any  $\lambda > 0$  and  $d \in \mathbb{N}$  we have*

$$\lim_{\varepsilon \downarrow 0} \frac{Y_d(\varepsilon)}{Y_d(\lambda\varepsilon)} = \lambda \quad \text{and} \quad \lim_{\varepsilon \downarrow 0} \frac{Y_{d-1}(\varepsilon)}{Y_d(\lambda\varepsilon)} = 0. \quad (4.24)$$

**Proof.** Clearly, for every  $0 < \delta < 1$  there exists  $0 < \varepsilon(\delta) < 1$  such that

$$1 \leq (\ln Y_d(\varepsilon))^{d-1} \leq Y_d(\varepsilon)^\delta \quad \text{for all } 0 < \varepsilon < \varepsilon(\delta),$$

and from (4.23) we obtain with  $c := 2^d/(d-1)!$  the inequality

$$c/2 \leq \frac{\varepsilon Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} \leq 2c.$$

This implies

$$c/2 \leq \frac{\varepsilon Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} \leq \varepsilon Y_d(\varepsilon) \quad \text{and} \quad 2c \geq \frac{\varepsilon Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} \geq \varepsilon Y_d(\varepsilon)^{1-\delta}.$$

Therefore

$$\frac{c}{2\varepsilon} \leq Y_d(\varepsilon) \leq \left(\frac{2c}{\varepsilon}\right)^{\frac{1}{1-\delta}}.$$

Hence we have, for all  $0 < \varepsilon < \varepsilon(\delta)$ ,

$$\frac{\ln(c/2) + \ln(1/\varepsilon)}{\ln(1/\varepsilon)} \leq \frac{\ln Y_d(\varepsilon)}{\ln(1/\varepsilon)} \leq \frac{\ln(2c) + \ln(1/\varepsilon)}{(1-\delta)\ln(1/\varepsilon)},$$

which implies

$$1 \leq \liminf_{\varepsilon \downarrow 0} \frac{\ln Y_d(\varepsilon)}{\ln(1/\varepsilon)} \leq \limsup_{\varepsilon \downarrow 0} \frac{\ln Y_d(\varepsilon)}{\ln(1/\varepsilon)} \leq \frac{1}{1-\delta}.$$

Since this is true for all  $0 < \delta < 1$ , we obtain  $\lim_{\varepsilon \downarrow 0} \ln Y_d(\varepsilon)/\ln(1/\varepsilon) = 1$ . In addition,

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon Y_d(\varepsilon)}{(\ln(1/\varepsilon))^{d-1}} = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon Y_d(\varepsilon)}{(\ln(Y_d(\varepsilon)))^{d-1}} \cdot \lim_{\varepsilon \downarrow 0} \left(\frac{\ln(Y_d(\varepsilon))}{\ln(1/\varepsilon)}\right)^{d-1} = \frac{2^d}{(d-1)!},$$

which finally gives

$$\lim_{\varepsilon \downarrow 0} \frac{Y_d(\varepsilon)}{Y_d(\lambda\varepsilon)} = \lambda \cdot \lim_{\varepsilon \downarrow 0} \frac{\varepsilon Y_d(\varepsilon)}{(\ln(1/\varepsilon))^{d-1}} \cdot \lim_{\varepsilon \downarrow 0} \frac{(\ln(1/(\lambda\varepsilon)))^{d-1}}{\lambda \varepsilon Y_d(\lambda\varepsilon)} \cdot \left(\lim_{\varepsilon \downarrow 0} \frac{\ln(1/\varepsilon)}{\ln(1/(\lambda\varepsilon))}\right)^{d-1} = \lambda.$$

The second identity in (4.24) follows with a similar argument. ■

**Lemma 4.14.** Let  $(z_\ell)_{\ell \in \mathbb{Z}}$  be a sequence indexed by  $\mathbb{Z}$  such that

$$0 < z_\ell \leq z_0 = 1 \quad \text{for all } \ell \neq 0 \quad \text{and} \quad \lim_{|\ell| \rightarrow \infty} \frac{y_\ell}{z_\ell} = 1.$$

Similarly as for  $(y_\ell)_{\ell \in \mathbb{Z}}$  we define  $\mathcal{Z}_d(\varepsilon)$  and  $Z_d(\varepsilon)$  associated to  $(z_\ell)_{\ell \in \mathbb{Z}}$ . Then we have

$$\lim_{\varepsilon \downarrow 0} \frac{Z_d(\varepsilon)}{Y_d(\varepsilon)} = 1. \quad (4.25)$$

**Proof.** Let us first observe that there are constants  $0 < c \leq C < \infty$  such that  $c \leq y_\ell/z_\ell \leq C$  for all  $\ell \in \mathbb{Z}$ . Fix now  $0 < \varepsilon \leq 1$  and  $\delta > 0$  (small), and select  $m = m(\delta) \in \mathbb{N}$  such that

$$1 - \delta \leq \frac{y_\ell}{z_\ell} \leq 1 + \delta \quad \text{for all } |\ell| \geq m.$$

For  $k \in \mathcal{Z}_d(\varepsilon)$ , we distinguish two cases.

Case 1,  $|k_j| \geq m$  for all  $j$ . This implies

$$\varepsilon \leq \prod_{j=1}^d z_{k_j} \leq \prod_{j=1}^d \frac{y_{k_j}}{1 - \delta},$$

and thus  $k \in \mathcal{Y}_d((1 - \delta)^d \varepsilon)$ .

Case 2,  $|k_\ell| < m$  for some  $\ell$ . Now we have

$$\varepsilon \leq \prod_{j=1}^d z_{k_j} = z_{k_\ell} \prod_{j \neq \ell} z_{k_j} \leq \prod_{j \neq \ell} \frac{y_{k_j}}{c},$$

which gives  $(k_1, \dots, k_{\ell-1}, k_{\ell+1}, \dots, k_d) \in \mathcal{Y}_{d-1}(c^{d-1} \varepsilon)$ . Since there are  $d$  choices of the index  $\ell \in \{1, \dots, d\}$  and  $2m - 1$  possible values of  $k_\ell$ , we conclude that

$$Z_d(\varepsilon) \leq Y_d((1 - \delta)^d \varepsilon) + (2m - 1)d \cdot Y_{d-1}(c^{d-1} \varepsilon).$$

Using the relations (4.24), this gives

$$\limsup_{\varepsilon \downarrow 0} \frac{Z_d(\varepsilon)}{Y_d(\varepsilon)} \leq \frac{1}{(1 - \delta)^d}. \quad (4.26)$$

Now we show a lower estimate for  $Z_d(\varepsilon)$ . Let  $k \in \mathcal{Y}_d((1 + \delta)^d \varepsilon)$ . Again we distinguish two cases.

If all  $|k_j| \geq m$ , we have

$$(1 + \delta)^d \varepsilon \leq \prod_{j=1}^d y_{k_j} \leq \prod_{j=1}^d (1 + \delta) z_{k_j},$$

that means  $k \in \mathcal{Z}_d(\varepsilon)$ .

Otherwise, if  $|k_\ell| < m$  for some  $\ell$ , we have

$$(1 + \delta)^d \varepsilon \leq y_{k_\ell} \cdot \prod_{j \neq \ell} y_{k_j} \leq \prod_{j \neq \ell} y_{k_j},$$

which means  $(k_1, \dots, k_{\ell-1}, k_{\ell+1}, \dots, k_d) \in \mathcal{Y}_{d-1}((1+\delta)^d \varepsilon)$ , and we get

$$Y_d((1+\delta)^d \varepsilon) - 2(m-1)d \cdot Y_{d-1}((1+\delta)^d \varepsilon) \leq Z_d(\varepsilon).$$

This implies, using again (4.24),

$$\liminf_{\varepsilon \downarrow 0} \frac{Z_d(\varepsilon)}{Y_d(\varepsilon)} \geq \frac{1}{(1+\delta)^d}, \quad (4.27)$$

and since (4.26) and (4.27) are true for all  $\delta > 0$ , the proof is finished.  $\blacksquare$

There are some simple consequences of Lemma 4.14 which are of interest for us. Taking logarithms in (4.25) yields

$$\lim_{\varepsilon \downarrow 0} \left( \ln Z_d(\varepsilon) - \ln Y_d(\varepsilon) \right) = 0.$$

Since  $\lim_{\varepsilon \downarrow 0} Y_d(\varepsilon) = \infty$ , we get

$$\lim_{\varepsilon \downarrow 0} \frac{\ln Z_d(\varepsilon)}{\ln Y_d(\varepsilon)} = \lim_{\varepsilon \downarrow 0} \frac{\ln Z_d(\varepsilon) - \ln Y_d(\varepsilon)}{\ln Y_d(\varepsilon)} + 1 = 1. \quad (4.28)$$

Clearly,

$$\frac{\varepsilon \cdot Z_d(\varepsilon)}{(\ln Z_d(\varepsilon))^{d-1}} = \frac{\varepsilon \cdot Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} \cdot \left( \frac{\ln Y_d(\varepsilon)}{\ln Z_d(\varepsilon)} \right)^{d-1} \cdot \frac{Z_d(\varepsilon)}{Y_d(\varepsilon)},$$

and together with (4.23), (4.25) and (4.28) this implies

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon \cdot Z_d(\varepsilon)}{(\ln Z_d(\varepsilon))^{d-1}} = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon \cdot Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} = \frac{2^d}{(d-1)!}. \quad (4.29)$$

### 4.3.3 The approximation numbers of $H_{\text{mix}}^{s,+}(\mathbb{T}^d)$ in $L_2(\mathbb{T}^d)$

In our first application of (4.23) we choose  $z_\ell := (1 + |\ell|^2)^{-1/2}$ ,  $\ell \in \mathbb{Z}$ . Then

$$\mathcal{Z}_d(\varepsilon) := \left\{ k \in \mathbb{Z}^d : \prod_{j=1}^d z_{k_j} \geq \varepsilon \right\} = \left\{ k \in \mathbb{Z}^d : 1/w_1^+(k) \geq \varepsilon \right\},$$

for all  $\varepsilon > 0$ .

**Corollary 4.15.** *Let  $d \in \mathbb{N}$ .*

(i) *Let  $s > 0$ . Then*

$$\lim_{n \rightarrow \infty} \frac{n^s a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{(d-1)s}} = \left[ \frac{2^d}{(d-1)!} \right]^s.$$

(ii) *Let  $s_0 > s_1 \geq 0$ . Then*

$$\lim_{n \rightarrow \infty} \frac{n^{s_0-s_1} a_n(I_d : H_{\text{mix}}^{s_0,+}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s_1,+}(\mathbb{T}^d))}{(\ln n)^{(d-1)(s_0-s_1)}} = \left[ \frac{2^d}{(d-1)!} \right]^{s_0-s_1}.$$



**Proof.** It is enough to prove (i) for  $s = 1$ . Indeed, then the known relation

$$a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = a_n(I_d : H_{\text{mix}}^{1,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))^s$$

implies (i) for arbitrary  $s > 0$ , and the semigroup property of the weights yields (ii).

Setting  $\varepsilon_r := (1 + r^2)^{-1/2}$  for  $r \in \mathbb{N}_0$ , we obviously have

$$\{\varepsilon_r : r \in \mathbb{N}\} \subset \{1/w_1^+(k) : k \in \mathbb{Z}^d\},$$

whence  $a_n := a_n(I_d : H_{\text{mix}}^{1,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = \varepsilon_r$  if  $n = Z_d(\varepsilon_r)$ . Consequently, if  $Z_d(\varepsilon_{r-1}) \leq n \leq Z_d(\varepsilon_r)$ , then

$$\varepsilon_r \leq a_n \leq \varepsilon_{r-1} \quad \text{and} \quad \frac{\varepsilon_r \cdot Z_d(\varepsilon_{r-1})}{(\ln Z_d(\varepsilon_{r-1}))^{d-1}} \leq \frac{na_n}{(\ln n)^{d-1}} \leq \frac{\varepsilon_{r-1} \cdot Z_d(\varepsilon_r)}{(\ln Z_d(\varepsilon_r))^{d-1}}.$$

Since  $\lim_{r \rightarrow \infty} \varepsilon_{r-1}/\varepsilon_r = 1$ , a simple monotonicity argument and (4.29) imply

$$\lim_{n \rightarrow \infty} \frac{na_n}{(\ln n)^{d-1}} = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon \cdot Z_d(\varepsilon)}{(\ln Z_d(\varepsilon))^{d-1}} = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon \cdot Y_d(\varepsilon)}{(\ln Y_d(\varepsilon))^{d-1}} = \frac{2^d}{(d-1)!}.$$

■

Corollary 4.15 is the basis for the two-sided estimates of  $a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for large  $n$  which we will study next.

**Theorem 4.16.** *Let  $s > 0$  and  $d \in \mathbb{N}$ . Then we have*

$$a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left[ \frac{(3 \cdot \sqrt{2})^d}{(d-1)!} \right]^s \frac{(\ln n)^{(d-1)s}}{n^s}, \quad \text{if } n \geq 27^d$$

and

$$a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq \left[ \frac{3}{d!} \left( \frac{2}{2 + \ln 12} \right)^d \right]^s \frac{(\ln n)^{(d-1)s}}{n^s}, \quad \text{if } n > (12e^2)^d.$$

**Proof.** By  $I^j$ ,  $j = 1, 2, 3$ , we denote identity mappings.

Since we have  $\|I^1 : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d)\| = 1$ , the commutative diagram

$$\begin{array}{ccc} H_{\text{mix}}^{s,\#}(\mathbb{T}^d) & \xrightarrow{I^1} & H_{\text{mix}}^{s,+}(\mathbb{T}^d) \\ & \searrow I^3 & \swarrow I^2 \\ & & L_2(\mathbb{T}^d) \end{array}$$

with  $I^3 = I^2 \circ I^1$  and basic properties of approximation numbers yields

$$a_n(I^3) \leq \|I^1 : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,+}(\mathbb{T}^d)\| a_n(I^2).$$

Supplemented by Theorem 4.5/(ii) the lower estimate of  $a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  follows. What concerns the upper bound we observe

$$\|I^1 : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,\#}(\mathbb{T}^d)\| = 2^{ds/2}$$

and consider the diagram

$$\begin{array}{ccc}
H_{\text{mix}}^{s,+}(\mathbb{T}^d) & \xrightarrow{I^1} & H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \\
& \searrow^{I^3} & \swarrow^{I^2} \\
& & L_2(\mathbb{T}^d)
\end{array}$$

with  $I^3 = I^2 \circ I^1$ . This leads to

$$a_n(I^3) \leq 2^{ds/2} a_n(I^2).$$

■

Finally, we shall have a look at the behavior of  $a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for small  $n$ . Recall that the quantity  $\alpha(n, d)$  has been defined in (4.15).

**Theorem 4.17.** *Let  $d \in \mathbb{N}$ ,  $d \geq 2$ , and  $s > 0$ . Then for any  $2 \leq n \leq \frac{d}{2}4^d$  it holds the two-sided estimate*

$$2^{-s} \left( \frac{1}{n} \right)^{\frac{s}{2+\alpha(n,d)}} \leq a_n(I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left( \frac{e^2}{n} \right)^{\frac{s}{4+2 \log_2 d}}.$$

**Proof.** The upper bound is a direct consequence of basic properties of approximation numbers, see Subsection 2.3, Lemma 2.2/(v), Theorem 4.9 and the first commutative diagram below, where  $I^3 = I^2 \circ I^1$ .

$$\begin{array}{ccccc}
H_{\text{mix}}^{s,+}(\mathbb{T}^d) & \xrightarrow{I^1} & H_{\text{mix}}^{s/2,\#}(\mathbb{T}^d) & H_{\text{mix}}^{s,\#}(\mathbb{T}^d) & \xrightarrow{I^4} & H_{\text{mix}}^{s,+}(\mathbb{T}^d) \\
& \searrow^{I^3} & \swarrow^{I^2} & \searrow^{I^5} & \swarrow^{I^3} & \\
& & L_2(\mathbb{T}^d) & & L_2(\mathbb{T}^d) &
\end{array}$$

The lower bound follows from Lemma 2.2/(i)-(iii), Theorem 4.10 and the second commutative diagram above, where  $I^5 = I^3 \circ I^4$ . ■

#### 4.3.4 The approximation numbers of $H_{\text{mix}}^{s,*}(\mathbb{T}^d)$ in $L_2(\mathbb{T}^d)$

In our second application of (4.23) we choose  $z_\ell := (1 + |\ell|^{2s})^{-1/(2s)}$ . This leads to

$$\mathcal{Z}_d(\varepsilon) := \left\{ k \in \mathbb{Z}^d : \prod_{j=1}^d z_{k_j} \geq \varepsilon \right\} = \left\{ k \in \mathbb{Z}^d : 1/w_s^*(k) \geq \varepsilon^s \right\}.$$

Due to the missing semigroup property we have to deal now with all  $s > 0$ , not only with  $s = 1$ . But nevertheless we can proceed similarly as in the previous subsection.

**Corollary 4.18.** *Let  $d \in \mathbb{N}$  and  $s > 0$ . Then*

$$\lim_{n \rightarrow \infty} \frac{n^s a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{(d-1)s}} = \left[ \frac{2^d}{(d-1)!} \right]^s.$$

**Proof.** Again (4.29) leads to

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon Z_d(\varepsilon)}{(\ln Z_d(\varepsilon))^{d-1}} = \frac{2^d}{(d-1)!}.$$

Setting  $\varepsilon_r := (1 + r^{2s})^{-1/2s}$ , we have  $\lim_{r \rightarrow \infty} \frac{\varepsilon_r}{\varepsilon_{r-1}}$ , and

$$\{\varepsilon_r^s : r \in \mathbb{N}_0\} \subset \{1/w_s^*(k) : k \in \mathbb{Z}^d\}.$$

Therefore,  $\varepsilon_r^s \leq a_n \leq \varepsilon_{r-1}$ , if  $Z_d(\varepsilon_r) \leq n \leq Z_d(\varepsilon_{r-1})$ . This gives

$$\varepsilon_r \leq a_n^{1/s} \leq \varepsilon_{r-1} \quad \text{and} \quad \frac{\varepsilon_r \cdot Z_d(\varepsilon_{r-1})}{(\ln Z_d(\varepsilon_{r-1}))^{d-1}} \leq \frac{na_n^{1/s}}{(\ln n)^{d-1}} \leq \frac{\varepsilon_{r-1} \cdot Z_d(\varepsilon_r)}{(\ln Z_d(\varepsilon_r))^{d-1}}.$$

Exactly the same arguments as in the proof of Theorem 4.15 imply

$$\lim_{n \rightarrow \infty} \frac{na_n^{1/s}}{(\ln n)^{d-1}} = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon \cdot Z_d(\varepsilon)}{(\ln Z_d(\varepsilon))^{d-1}} = \frac{2^d}{(d-1)!},$$

and this is equivalent to our assertion. ■

Based on Corollary 4.18 we can derive two-sided estimates of  $a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for large  $n$ .

**Theorem 4.19.** *Let  $d \in \mathbb{N}$ .*

(i) *Let  $s > 1/2$ . Then*

$$a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq 2^{-d/2} \left[ \frac{6^d}{(d-1)!} \right]^s \frac{(\ln n)^{(d-1)s}}{n^s} \quad \text{if } n \geq 27^d,$$

and

$$a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq \left[ \frac{3 \cdot 2^d}{d!(2 + \ln 12)^d} \right]^s \frac{(\ln n)^{(d-1)s}}{n^s} \quad \text{if } n > (12e^2)^d.$$

(ii) *Let  $0 < s \leq 1/2$ . Then*

$$a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left[ \frac{3^d}{(d-1)!} \right]^s \frac{(\ln n)^{(d-1)s}}{n^s} \quad \text{if } n \geq 27^d,$$

and

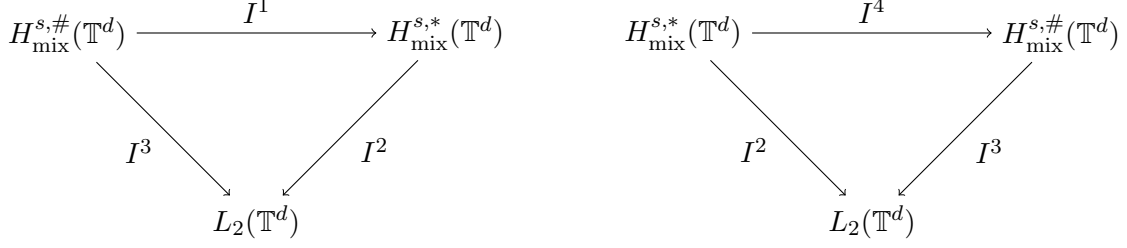
$$a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq 2^{-d/2} \left[ \frac{3 \cdot 4^d}{d!(2 + \ln 12)^d} \right]^s \frac{(\ln n)^{(d-1)s}}{n^s} \quad \text{if } n > (12e^2)^d.$$

**Proof.** We distinguish two cases:  $s > 1/2$  and  $0 < s \leq 1/2$ .

*Step 1.* Let  $s > 1/2$ . Then

$$\|I^1 : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,*}(\mathbb{T}^d)\| = 1 \quad \text{and} \quad \|I^4 : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,\#}(\mathbb{T}^d)\| = 2^{(s-1/2)d}.$$

In view of the diagrams



with  $I^3 = I^2 \circ I^1$ ,  $I^2 = I^3 \circ I^4$ , this yields

$$a_n(I^3) \leq a_n(I^2) \leq 2^{(s-1/2)d} a_n(I^3).$$

Now the claimed estimates follow from Theorem 4.5.

*Step 2.* Let  $0 < s \leq 1/2$ . Then

$$\|I^1 : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,*}(\mathbb{T}^d)\| = 2^{(\frac{1}{2}-s)d} \quad \text{and} \quad \|I^4 : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow H_{\text{mix}}^{s,\#}(\mathbb{T}^d)\| = 1.$$

Using the same diagrams as in Step 1 we conclude

$$a_n(I^2) \leq a_n(I^3) \leq 2^{(\frac{1}{2}-s)d} a_n(I^2).$$

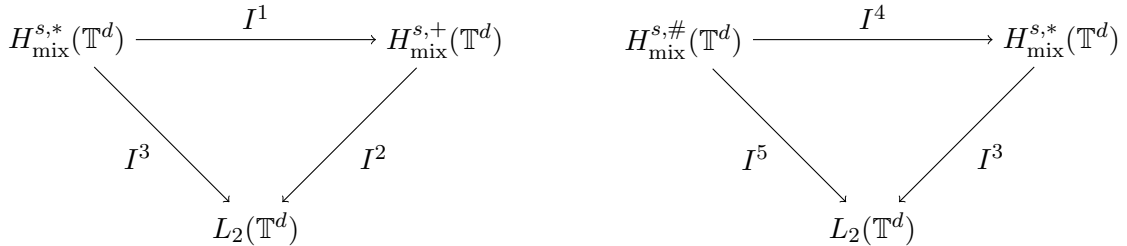
Now the claimed estimates follow from Theorem 4.5. ■

Again, as the last step in this subsection, we shall consider the behavior of  $a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for small  $n$ . This time we have only partial results.

**Theorem 4.20.** *Let  $d \in \mathbb{N}$ ,  $d \geq 2$ , and  $1/2 \leq s \leq 1$ . Then for any  $2 \leq n \leq \frac{d}{2}4^d$  it holds the two-sided estimate*

$$2^{-s} \left(\frac{1}{n}\right)^{\frac{s}{2+\alpha(n,d)}} \leq a_n(I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left(\frac{e^2}{n}\right)^{\frac{s}{4+\log_2(d^2)}}.$$

**Proof.** We argue as in the proof of Theorem 4.17. The upper bound is a consequence of Lemma 2.2,(ii) together with Theorem 4.17 and the first commutative diagram



with  $I^3 = I^2 \circ I^1$ . The lower bound follows from Lemma 2.2, (ii), Theorem 4.10 and the second commutative diagram using  $I^5 = I^3 \circ I^4$ . ■

#### 4.4 The approximation numbers of $H_{\text{mix}}^m(\mathbb{T}^d)$ in $L_2(\mathbb{T}^d)$

By setting  $z_\ell = v_m(\ell)^{1/m}$ , see (4.22), we could argue similar as in the previous subsection to compute  $\lim_{n \rightarrow \infty} n^m \cdot a_n / (\ln n)^{(d-1)m}$ . However, Lemma 2.3 provides a much simpler argument.

**Corollary 4.21.** *Let  $d \in \mathbb{N}$  and  $m \in \mathbb{N}$ . Then*

$$\lim_{n \rightarrow \infty} \frac{n^m a_n(I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))}{(\ln n)^{(d-1)m}} = \left[ \frac{2^d}{(d-1)!} \right]^m.$$

**Proof.** This follows immediately from Corollaries 4.15 and 4.18, and

$$\begin{aligned} a_n(I_d : H_{\text{mix}}^{m,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) &\leq a_n(I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \\ &\leq a_n(I_d : H_{\text{mix}}^{m,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \end{aligned} \quad (4.30)$$

which is itself a consequence of Lemma 2.3. ■

Based on Theorems 4.16, 4.19 and (4.30) we derive two-sided estimates of  $a_n(I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for large  $n$ .

**Theorem 4.22.** *Let  $m \in \mathbb{N}$  and  $d \in \mathbb{N}$ .*

(i) *Then*

$$a_n(I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left[ \frac{6^d}{(d-1)!} \right]^m 2^{-d/2} \frac{(\ln n)^{(d-1)m}}{n^m} \quad \text{if } n \geq 27^d.$$

(ii) *In addition*

$$a_n(I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \geq \left[ \frac{3 \cdot 2^d}{d! (2 + \ln 12)^d} \right]^s \frac{(\ln n)^{(d-1)m}}{n^m} \quad \text{if } n > (12 e^2)^d.$$

**Remark 4.23.** Also for the embedding  $H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$  the behavior of  $a_n(I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  for small  $n$  is of interest. By the coincidence  $H_{\text{mix}}^1(\mathbb{T}^d) = H_{\text{mix}}^{1,*}(\mathbb{T}^d)$  (with equal norms) we obtain the relations

$$\frac{1}{2} \left( \frac{1}{n} \right)^{\frac{1}{2+\alpha(n,d)}} \leq a_n(I_d : H_{\text{mix}}^1(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left( \frac{e^2}{n} \right)^{\frac{1}{4+\log_2(d^2)}}$$

immediately from Theorem 4.20.

## 4.5 Various comments on the literature

- We comment on (1.2). The first who considered this problem was Babenko [1], but formula (1.2) cannot be found in [1]. Babenko gave a two-sided estimate of the  $a_n$  in terms of the eigenvalues of an associated operator. It seems that (1.2) is due to Mityagin [17]. In his proof he used partial Fourier sums with respect to hyperbolic crosses. In fact, Mityagin studied the behaviour of the Kolmogorov  $n$ -widths  $d_n(I_d : S_p^r W(\mathbb{T}^d) \rightarrow L_p(\mathbb{T}^d))$  for arbitrary  $1 < p < \infty$ . Recall that  $a_n(I_d : S_2^r W(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = d_n(I_d : S_2^r W(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$  and  $H_{\text{mix}}^s(\mathbb{T}^d) = S_2^s W(\mathbb{T}^d)$ . The classes  $S_p^r W(\mathbb{T}^d)$  denote the periodic Sobolev spaces of dominating mixed smoothness  $r$  and integrability  $p$ , where  $1 < p < \infty$ . Later on, Bugrov [2] applied similar methods for obtaining upper bounds for  $a_n(I_d : S_{p,\infty}^r B(\mathbb{T}^d) \rightarrow L_p(\mathbb{T}^d))$ , where  $S_{p,\infty}^r B(\mathbb{T}^d)$  denotes the related periodic Nikol'skij-Besov space of dominating mixed smoothness. These investigations from the early 1960s have been continued by Nikol'skaya [18], Galeev [9], Temlyakov [32], Dinh Dũng [6] and many others. However, all those authors did not consider the  $d$ -dependence of the constants.

- Closest to us in aims and methods is the recent paper [8]. There, in Theorem 3.13, the authors obtained for  $s > 0$  and any  $n \geq 2^d$  the inequality

$$a_n(I_d : H_{\text{mix}}^{s, \square}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq 4^s \left( \frac{2e}{d-1} \right)^{s(d-1)} n^{-s} (\log_2 n)^{s(d-1)}, \quad (4.31)$$

where  $\square$  indicates that the space  $H_{\text{mix}}^{s, \square}(\mathbb{T}^d)$  is equipped with a further norm (based on dyadic decompositions on the Fourier side and different from those studied here). This has to be compared with Theorems 4.5, 4.16, 4.19 and 4.22. In all these cases we have a super-exponential decay of the constants  $C_s(d)$  in  $d$ .

- Super-exponential decay of the constants  $C_s(d)$  in  $d$  has been observed before. Bungartz and Griebel [4, Theorem 3.8] investigated the non-periodic situation. An approximation is given with respect to tensor products of piecewise linear functions. The authors proved that for any  $n \in \mathbb{N}$  there exists a subspace  $V_n^{(1)} \subset L_2([0, 1]^d)$  with  $m = m(d, n)$  degrees of freedom and a projection  $Q_n$  onto  $V_n^{(1)}$  such that

$$\|f - Q_n f\|_{L_2([0, 1]^d)} \leq \frac{2}{12^d} 2^{-2n} A(d, n) \left\| \frac{\partial^{2d} f}{\partial x_1^2 \dots \partial x_d^2} \Big|_{L_2([0, 1]^d)} \right\| \quad (4.32)$$

holds for all continuous functions  $f$  vanishing on the boundary  $\partial([0, 1]^d)$ . The latter assumption is actually crucial for the explicit bound in (4.32). Here, the number  $A(d, n)$  is given by

$$A(d, n) := \sum_{k=0}^{d-1} \binom{n+d-1}{k}.$$

Of course, inequality (4.32) does not allow for a comparison with the results obtained in this paper. One first has to rewrite the bounds in terms of the degrees of freedom  $m$ . For this issue a careful two-sided estimate of  $m = \dim V_n^{(1)}$  is required first. Lemma 3.6 in [4] shows that  $\dim V_n^{(1)}$  can be estimated from above and below by

$$2^{n-1} \binom{n+d-2}{d-1} \leq \dim V_n^{(1)} = \sum_{j=0}^{n-1} 2^j \binom{j+d-1}{d-1} \leq 2^n \binom{n+d-2}{d-1}. \quad (4.33)$$

For  $n \geq d-1$  the expression  $\binom{n+d-1}{k}$  is increasing in  $k \leq d-1$ . In this case we have the estimate

$$A(d, n) \leq d \binom{n+d-1}{d-1}.$$

We will now transfer (4.32) to the notion of approximation numbers. To be precise we consider the space/norm

$$H_{\text{mix},0}^2([0, 1]^d) := \{f \in L_2([0, 1]^d) : \|f\|_{H_{\text{mix}}^2([0, 1]^d)} := \max_{\alpha \in \{0,2\}^d} \|D^\alpha f\|_{L_2([0, 1]^d)} < \infty \\ \text{and } f = 0 \text{ on the boundary}\}.$$

Note, that  $\|f\|_{H_{\text{mix}}^2([0, 1]^d)}$  in the previous formula represents a very weak norm in  $H_{\text{mix},0}^2([0, 1]^d)$  compared to (2.2). Based on (4.32), (4.33) and the monotonicity of approximation numbers we find for any  $m$  satisfying  $\dim V_n^{(1)} < m \leq \dim V_{n+1}^{(1)}$  for some

$n \geq d - 1$  the relation

$$\begin{aligned} a_m(I_d : H_{\text{mix},0}^2([0,1]^d) \rightarrow L_2([0,1]^d)) &\leq \frac{2}{12^d} 2^{-2n} A(d,n) \leq \frac{8d}{12^d} m^{-2} \binom{n+d-1}{d-1}^3 \\ &\leq \frac{8d}{12^d} m^{-2} \binom{\log_2 m + d}{d-1}^3 \\ &\leq \frac{8d}{12^d} m^{-2} \left( \frac{2e \log_2 m}{d-1} \right)^{3(d-1)}. \end{aligned}$$

Consequently, inequality (4.32) implies for any  $m \geq 2^{d-1} \binom{2d-2}{d-1}$  that

$$a_m(I_d : H_{\text{mix},0}^2([0,1]^d) \rightarrow L_2([0,1]^d)) \leq C(d) m^{-2} (\log_2 m)^{3(d-1)}, \quad (4.34)$$

where

$$C(d) := \frac{2d}{3} \left( \frac{2e}{12^{1/3}(d-1)} \right)^{3(d-1)}.$$

This constant  $C(d)$  is decaying extremely fast in  $d$ , i.e. super-exponentially, similarly as in Theorems 4.5, 4.16, 4.19, 4.22 or in (4.31) above. But comparing (4.34) with the quoted estimates of the approximation numbers  $a_n(I_d : H_{\text{mix}}^2(\mathbb{T}^d), L_2(\mathbb{T}^d))$ , then it is obvious that the power of the logarithm  $3(d-1)$  in (4.34) is larger than there, where it is always  $2(d-1)$ . This is at least partly caused by the fact that interpolation operators of Smolyak type are known to be not optimal in the sense of approximation numbers in such a context and Bungartz and Griebel are using an interpolation operator of Smolyak type with respect to a sparse grid. However, one cannot really compare (4.32) with Theorems 4.5, 4.16, 4.19, 4.22 or (4.31), due to the following reasons:

- (i) The periodic Sobolev spaces  $H_{\text{mix}}^2(\mathbb{T}^d)$  are smaller than the non-periodic Sobolev spaces  $H_{\text{mix}}^2([0,1]^d)$  (and the “difference” is increasing with  $d$ ).
  - (ii) The space  $H_{\text{mix},0}^2([0,1]^d)$  is much smaller than the original space  $H_{\text{mix}}^2([0,1]^d)$ .
  - (iii) On the right-hand side in (4.32) only the term  $\left\| \frac{\partial^{2d} f}{\partial x_1^2 \dots \partial x_d^2} \Big|_{L_2([0,1]^d)} \right\|$  shows up, which is much smaller than the full norm used in our investigations above.
- Preasymptotics. The inequalities (4.32) and (4.33) remain also true for small  $n$ . Note, that in case  $1 \leq n \leq d-1$  the number  $A(d,n)$  can be estimated as follows

$$\frac{1}{2} 2^{n+d-1} \leq A(d,n) = \sum_{k=0}^{d-1} \binom{n+d-1}{k} \leq 2^{n+d-1}$$

(we sum up to  $d-1$ , which is larger than  $(n+d-1)/2$ ). Let  $\dim V_n^{(1)} \leq m \leq \dim V_{n+1}^{(1)}$  for some  $1 \leq n \leq d-1$ . Using (4.33), (4.32) and the space  $H_{\text{mix},0}^2([0,1]^d)$  defined above we obtain

$$\begin{aligned} a_m(I_d : H_{\text{mix},0}^2([0,1]^d) \rightarrow L_2([0,1]^d)) &\leq \frac{2}{12^d} 2^{-2n} A(d,n) \\ &\leq \frac{4}{12^d} m^{-1} 2^{n+d-1} 2^{-n} \binom{n+d-1}{d-1} \\ &\leq \frac{1}{3} \left( \frac{4e}{12} \right)^{d-1} m^{-1} = \frac{1}{e} \left( \frac{e}{3} \right)^d m^{-1}. \end{aligned}$$

Hence we obtain for  $1 \leq m \leq 2^{d-1} \binom{2d-2}{d-1}$  the ‘‘preasymptotic’’ decay

$$a_m(I_d : H_{\text{mix},0}^2([0,1]^d) \rightarrow L_2([0,1]^d)) \leq C(d)m^{-1} \quad (4.35)$$

with

$$C(d) := \frac{1}{e} \left(\frac{e}{3}\right)^d.$$

This time we just have exponential decay of the constant  $C(d)$  in  $d$ . Now we compare this with our results obtained in Theorems 4.9, 4.17 and 4.20. Let us concentrate on formula (4.12), which implies for  $s = 2$  and  $1 \leq m \leq 4^d$  the inequality

$$a_m(I_d : H_{\text{mix}}^{2,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) \leq \left(\frac{e^2}{m}\right)^{\frac{1}{1+\log_2 \sqrt{d}}}.$$

Inequality (4.35) looks much better than this inequality (with respect to the exponent of  $m$  as well as with respect to the  $d$ -dependence of the constant). This is mainly caused by the homogeneous boundary condition in (4.35) which shrinks the space significantly.

- Also Schwab, Süli, and Todor [26] considered the non-periodic situation. A particular case of their results in [26, Theorem 5.1] can be formulated as follows. For  $s \in \mathbb{N}$ ,  $s \geq 2$  and  $L \geq d - 1$  there exists a subspace  $V_L$  of  $L_2([0,1]^d)$  with  $m = \dim V_L$  degrees of freedom and a projection  $Q_L$  onto  $V_L$  such that

$$\|f - Q_L f\|_{L_2([0,1]^d)} \leq C_2(d, s) m^{-s} (\log m)^{(d-1)(s+1)} \|f\|_{H_{\text{mix}}^{s,\Delta}([0,1]^d)}, \quad (4.36)$$

where  $C_2(d, s)$  decays super-exponentially in  $d$ . Here the authors use the stronger norm

$$\|f\|_{H_{\text{mix}}^{s,\Delta}([0,1]^d)} := \sum_{\substack{0 \leq \alpha_i \leq s \\ i=1,\dots,d}} \|D^\alpha f\|_{L_2([0,1]^d)},$$

where  $s \in \mathbb{N}$ . The result is stated in [26] in a slightly different form. As above one has to supplement an inequality similar to (4.32) by two-sided estimates for  $\dim V_L$  to turn it into (4.36). In some sense (4.36) generalizes (4.32) to the case of higher smoothness. As before the power of the logarithm is worse compared with the behavior of the approximation numbers  $a_m(I_d : H_{\text{mix}}^s(\mathbb{T}^d), L_2(\mathbb{T}^d))$ .

- Preasymptotics. In case  $L \leq d-1$  Schwab, Süli, and Todor [26] obtained, under additional restrictions, the estimate

$$\|f - Q_L f\|_{L_2([0,1]^d)} \leq C_3(d, s) 2^{-Ls} \|f\|_{H_{\text{mix}}^{s,\Delta}([0,1]^d)},$$

where  $C_3(d, s)$  decays exponentially in  $d$ . Again this has to be complemented by a two-sided estimate of  $m := \dim V_L$ . A rather rough estimate yields

$$\|f - Q_L f\|_{L_2([0,1]^d)} \leq C_4(d, s) m^{-s/3} \|f\|_{H_{\text{mix}}^{s,\Delta}([0,1]^d)}, \quad m \leq 2^{2(d-1)},$$

where we have been unable to clarify the dependence of the constant  $C_4(d, s)$  on  $d$ .

- Neither Bungartz and Griebel [4] nor Schwab, Süli, and Todor [26] considered estimates from below.



- Sampling operators versus general linear operators. As already mentioned above, the estimates (4.34) and (4.36) have been obtained by using interpolation operators  $Q_n$  with  $m(d, n)$  sample points based upon univariate spline interpolation operators (via a Smolyak construction). Let us mention the following result in this context: In [30, Theorem 6.2,(i)] two of the present authors showed for  $1/2 < s < 2$  and all  $m \in \mathbb{N}$

$$\|f - A_m f\|_{L_2([0, 1]^d)} \lesssim m^{-s} (\log m)^{(d-1)(s+1/2)} \|f\|_{H_{\text{mix}}^s([0, 1]^d)}. \quad (4.37)$$

Compared to the estimates (4.34) and (4.36) this improves the power of the logarithm by  $(d-1)/2$ , however, we do not know about the  $d$ -dependence of the constants in (4.37). The restrictions on  $s$  in (4.37) are caused by the fact that we worked with piecewise linear functions. Dinh Dũng [7] extended the relation (4.37) to all  $s > 1/2$  via  $B$ -spline quasi-interpolation, but also without taking care of the  $d$ -dependence of the constants.

- Motivated by the aim to approximate the solution of a Poisson equation in the energy norm, i.e., in the norm of the isotropic Sobolev space  $H^1$ , Bungartz and Griebel [3] investigated estimates of the quantities  $a_n(I_d : H_{\text{mix}}^2([0, 1]^d), H^1([0, 1]^d))$ . These studies have been continued in Griebel, Knapek [12, 13], Bungartz, Griebel [4], Griebel [11], Schwab, Süli, Todor [26], and Dinh Dũng, Ullrich [8]. Let us comment on the non-periodic situation first. It was already noticed by Griebel in [11, Theorem 2] that in this situation the constant (in front of the approximation order term) decays exponentially in  $d$ . To be more precise, he proved that there is a subspace  $V_n$  with  $n$  degrees of freedom and a projection  $Q_n$  onto  $V_n$  such that for large  $n$

$$\|f - Q_n f\|_{H^1([0, 1]^d)} \leq c \cdot c_1(d) \cdot c_2(d) \cdot n^{-1} \left\| \frac{\partial^{2d} f}{\partial x_1^2 \dots \partial x_d^2} \Big|_{L_\infty[0, 1]^d} \right\|, \quad (4.38)$$

holds, where

$$c_1(d) = \frac{d}{2} e^d \quad \text{and} \quad c_2(d) = \frac{d}{3^{(d-1)/2} 4^{d-1}} \left[ \frac{1}{2} + \left( \frac{5}{2} \right)^{d-1} \right].$$

Hence, the product  $c_1(d)c_2(d)$  decays like  $d^2 \cdot 0.980875^d$ . Note, that the  $L_\infty$ -norm is involved in (4.38) and the functions  $f$  are taken from spaces with mixed smoothness of order 2 and homogeneous boundary conditions. The situation changes significantly if one replaces  $L_\infty$  by  $L_2$  in (4.38). The source space for  $f$  is now getting larger and hence the approximation is getting worse. In [4, Table 3.2, page 35] the constant  $c_2(d)$  can be chosen as

$$c_2(d) = \frac{2d}{\sqrt{3} \cdot 6^{d-1}} \left[ \frac{1}{2} + \left( \frac{5}{2} \right)^{d-1} \right].$$

Therefore,  $c_1(d)c_2(d)$  can only be estimated by  $d^2 \cdot 1.1326^d$  and thus an exponential decay cannot longer be guaranteed. However, if the smoothness  $s$  of the source space is less than 2 we can say something in the periodic setting. If the error is measured in  $H^1(\mathbb{T}^d)$  and  $s < 2$  it follows from [8, Theorem 3.6,(ii)] that for  $n > \lambda^d$  (for some  $\lambda > 1$ ) the relation

$$a_n(I_d : \tilde{H}_{\text{mix}}^s(\mathbb{T}^d) \rightarrow H^1(\mathbb{T}^d)) \leq c d^{s-1} \left( \frac{1}{2^{1/(s-1)} - 1} \right)^d n^{-(s-1)}$$

holds true. Here  $\tilde{H}_{\text{mix}}^s(\mathbb{T}^d)$  is the subspace of  $H_{\text{mix}}^s(\mathbb{T}^d)$  containing all functions  $f$  such that  $c_k(f) \neq 0 \implies \prod_{i=1}^d k_i \neq 0$ . This technical condition is essential to prove (4.38),

see [11, Theorem 2]. Without this condition, i.e. for the entire space  $H_{\text{mix}}^s(\mathbb{T}^d)$  normed with dyadic building blocks on the Fourier side, there is no exponential decay of the constants if  $s \leq 2$ , see [8, Theorem 4.7,(i)].

## 5 Quasi-polynomial tractability

Now we translate our results into tractability notions. Various concepts of tractability are discussed in the recent monographs by Novak and Woźniakowski [19, 21, 22]. We will obtain “quasi-polynomial tractability” of the respective approximation problems. This notion has been recently introduced in [10] and is a stronger notion than “weak tractability”.

### 5.1 General notions of tractability

For arbitrary  $s > 0$  and all  $d \in \mathbb{N}$  we consider the embedding operators (formal identities)

$$I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d),$$

where the Sobolev spaces are equipped with the norms  $\|\cdot\|_{H_{\text{mix}}^s(\mathbb{T}^d)}$ ,  $\|\cdot\|_{|H_{\text{mix}}^s(\mathbb{T}^d)|^*}$ , and  $\|\cdot\|_{|H^s(\mathbb{T}^d)|^\#}$ . In all three cases we have  $\|I_d\| = 1$  for all  $s > 0$  and  $d \in \mathbb{N}$ . In other words, the *normalized error criterion* is satisfied. In this context, a *linear algorithm* that uses *arbitrary information* is of the form

$$A_{n,d}(f) = \sum_{j=1}^n L_j(f)g_j,$$

where  $g_j \in L_2(\mathbb{T}^d)$  and  $L_j$  are continuous linear functionals. If the error is measured in the norm of  $L_2(\mathbb{T}^d)$  we can identify the algorithm  $A_{n,d}$  with a bounded linear operator  $A_{n,d} : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$  of rank at most  $n$ . The worst-case error of  $A_{n,d}$  with respect to the unit ball (respective norms) in  $H_{\text{mix}}^s(\mathbb{T}^d)$

$$\sup_{\|f\|_{H_{\text{mix}}^s(\mathbb{T}^d)} \leq 1} \|f - A_{n,d}(f)\|_{L_2(\mathbb{T}^d)}$$

clearly coincides with the operator norm  $\|I_d - A_{n,d} : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)\|$ , and the  *$n$ th minimal worst-case error* with respect to *linear algorithms and general information*

$$\inf_{\text{rank } A_{n,d} \leq n} \|I - A_{n,d} : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L(\mathbb{T}^d)\|$$

is just the *approximation number*  $a_{n+1}(I_d : H_{\text{mix}}^s(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d))$ , see (1.1).

Finally, the *information complexity* of the  $d$ -variate approximation problem is measured by the quantity  $n(\varepsilon, d)$  defined for  $\varepsilon > 0$  and  $d \in \mathbb{N}$  by

$$n(\varepsilon, d) := \inf\{n \in \mathbb{N} : a_n(I_d) \leq \varepsilon\}.$$

The approximation problem is called *weakly tractable*, if

$$\lim_{1/\varepsilon + d \rightarrow \infty} \frac{\ln n(\varepsilon, d)}{1/\varepsilon + d} = 0, \tag{5.1}$$

i.e.,  $n(\varepsilon, d)$  neither depends exponentially on  $1/\varepsilon$  nor on  $d$ . The problem is called *intractable*, if (5.1) does not hold, see the definition [19, p. 7]. If for some  $0 < \varepsilon < 1$  the number  $n(\varepsilon, d)$

is an exponential function in  $d$ , then we say that the approximation problem suffers from *the curse of dimensionality*. In other words, if there exist positive numbers  $c, \varepsilon_0, \gamma$  such that

$$n(\varepsilon, d) \geq c(1 + \gamma)^d, \quad \text{for all } 0 < \varepsilon \leq \varepsilon_0 \text{ and infinitely many } d \in \mathbb{N},$$

then the problem suffers from *the curse of dimensionality*.

We need a further notion of tractability, namely *quasi-polynomial tractability*, see for instance [10]. The approximation problem is called quasi-polynomially tractable if there are positive numbers  $t$  and  $C_t$  such that

$$n(\varepsilon, d) \leq C_t \exp(t \ln(\varepsilon^{-1})(1 + \ln(d))). \quad (5.2)$$

Of course, quasi-polynomial tractability implies weak tractability.

## 5.2 Tractability results for $H_{\text{mix}}^s(\mathbb{T}^d)$

By our results in Section 4 we are very well prepared for the investigation of these tractability problems, resulting in short proofs of the assertions.

**Theorem 5.1.** *For every  $s > 0$  the approximation problem for the embeddings*

$$I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$$

*is quasi-polynomially tractable.*

**Proof.** Let  $0 < \varepsilon \leq 1$  be given. Select  $r \in \mathbb{N}$  such that  $r^{-s} < \varepsilon \leq (r-1)^{-s}$ . Applying Lemma 4.1 we get  $a_{C(r,d)}(I_d : H_{\text{mix}}^{s, \#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)) = 1/r^s < \varepsilon$ . This implies

$$n(\varepsilon, d) \leq n(r^{-s}, d) \leq C(r, d).$$

Once again, the problem reduces to the estimation of  $C(r, d)$ . Parts of it have already been done in the proof of Theorem 4.9 above. There we observed that  $C(r, d) \leq e^2 r^{2 + \log_2 d}$  if  $r \leq 2^d$ . In case  $r \geq e^d$  we proved  $C(r, d) \leq r^2 3^d \leq r^4$ , see (4.7). It remains to deal with  $2^d < r < e^d$ . Based on Lemma 3.1, Lemma 3.2 and taking into account  $x^k/k! \leq e^x$  applied to  $x = \ln r$  we conclude

$$C(r, d) \leq 1 + \sum_{\ell=1}^d \binom{d}{\ell} 2^\ell r \frac{(\ln r)^{\ell-1}}{(\ell-1)!} \leq r^2 \sum_{\ell=1}^d \binom{d}{\ell} 2^\ell \leq 3^d r^2 \leq r^4.$$

This gives

$$\ln(n(\varepsilon, d)) \leq \ln C(r, d) \leq \begin{cases} 2 + (2 + \log_2 d) \ln r & : r \leq 2^d, \\ 4 \ln r & : r \geq 2^d. \end{cases}$$

Since  $\ln r \sim \ln(1/\varepsilon)/s$  we obtain (5.2). The proof is complete. ■

**Corollary 5.2.** *For every  $s > 0$  and every  $m \in \mathbb{N}$  the approximation problems for the embeddings*

$$I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d) \quad , \quad I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d) \quad \text{and} \quad I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$$

*are quasi-polynomially tractable.*

**Proof.** The results are direct consequences of the embeddings in Lemma 2.2. Because of Lemma 2.2/(iv) it is sufficient to consider  $s < 1/2$ . Then Lemma 2.2/(iii) and Theorem 5.1 immediately imply the statement for  $H_{\text{mix}}^{s,*}(\mathbb{T}^d)$ . Finally, Lemma 2.2/(v) and Theorem 5.1 give quasi-polynomial tractability for  $H_{\text{mix}}^{s,+}(\mathbb{T}^d)$ . ■

**Remark 5.3.** Tensor product problems play an essential role in IBC (information based complexity), see, e.g., Section 2 in Chapter 5 and Section 2 in Chapter 8 of the monograph [19]. The spaces  $H_{\text{mix}}^m(\mathbb{T}^d)$ ,  $H_{\text{mix}}^{s,\#}(\mathbb{T}^d)$ ,  $H_{\text{mix}}^{s,+}(\mathbb{T}^d)$ , and  $H_{\text{mix}}^{s,*}(\mathbb{T}^d)$  are  $d$ -fold tensor products of the univariate spaces  $H^m(\mathbb{T})$ ,  $H^{s,\#}(\mathbb{T})$ ,  $H^{s,+}(\mathbb{T})$ , and  $H^{s,*}(\mathbb{T})$ , respectively, see for instance [29]. Obviously, the identity  $I_d$  is a compact tensor product operator (considered as a mapping into the tensor product space  $L_2(\mathbb{T}^d)$ ). Since the approximation numbers decay polynomially in these four univariate situations and  $1 = a_1 > a_2$ , we obtain the following conclusion from the general Theorem 3.3 of [10]: For any  $s > 0$ , all four problems  $I_d : H_{\text{mix}}^m(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$ ,  $I_d : H_{\text{mix}}^{s,\#}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$ ,  $I_d : H_{\text{mix}}^{s,+}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$ , and  $I_d : H_{\text{mix}}^{s,*}(\mathbb{T}^d) \rightarrow L_2(\mathbb{T}^d)$  are quasi-polynomially tractable (and polynomially intractable). Let us also mention, that this result in [10] has a forerunner in [20], where it has been proven that such tensor product problems are weakly tractable. Hence, Theorem 5.1 and Corollary 5.2 are special cases of a more general theory. However, the approach given here is different.

**Acknowledgments.** The authors would like to thank the organizers of the Dagstuhl seminar 12391 “Algorithms and Complexity for Continuous Problems”, 2012, and the organizers of the HCM workshop “Discrepancy, Numerical Integration, and Hyperbolic Cross Approximation” where this work has been initiated and pursued, for providing a pleasant and fruitful working atmosphere. They would also like thank Dinh Dũng for commenting on earlier versions of this manuscript and pointing out the connection to [5] in Remark 3.3. Last but not least, the authors would like to thank an anonymous referee for asking for the details in Lemma 4.13.

## References

- [1] K.I. BABENKO, About the approximation of periodic functions of many variable trigonometric polynomials, *Dokl. Akad. Nauk SSR* **32** (1960), 247–250.
- [2] YA.S. BUGROV, Approximation of a class of functions with a dominant mixed derivative, *Mat. Sbornik* **64** (106) (1964), 410–418.
- [3] H.-J. BUNGARTZ AND M. GRIEBEL, A note on the complexity of solving Poisson’s equation for spaces of bounded mixed derivatives, *J. Complexity* **15** (1999), 167–199.
- [4] H.-J. BUNGARTZ AND M. GRIEBEL, Sparse grids, *Acta Numer.* **13** (2004), 147–269.
- [5] A. CHERNOV AND D. DŨNG, New explicit-in-dimension estimates for the cardinality of high-dimensional hyperbolic crosses and approximation of functions having mixed smoothness, ArXiv e-prints, 2014, arXiv:1309.5170 [math.NA].
- [6] D. DŨNG, Approximation of classes of functions according to the theory given by mixed modulus of continuity. In: *Constructive theory of functions*, Publishing House of the Bulgarian Academy of Sciences, Sofia, 1984, 43-48.

- [7] D. DŨNG, B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness, *Journal of Complexity* **27** (2011), 541–567.
- [8] D. DŨNG AND T. ULLRICH,  $N$ -Widths and  $\varepsilon$ -dimensions for high-dimensional approximations, *Found. Comp. Math* **13** (2013), 965–1003.
- [9] E.M. GALEEV, Approximation by Fourier sums of classes of functions with several bounded derivatives, *Mat. Zametki* **23** (1978), 197–212.
- [10] M. GNEWUCH AND H. WOŹNIAKOWSKI, Quasi-polynomial tractability, *J. Complexity* **27** (3-4) (2011), 312–330.
- [11] M. GRIEBEL, Sparse grids and related approximation schemes for higher dimensional problems. Proc. Foundations of Comp. Math., Santander 2005, 106–161, London Math. Soc. Lect. Notes Series, **331**, Cambridge Univ. Press, Cambridge, 2006.
- [12] M. GRIEBEL AND S. KNAPEK, Optimized tensor-product approximation spaces, *Constructive Approximation* **16**(4) (2000), 525–540.
- [13] M. GRIEBEL AND S. KNAPEK, Optimized general sparse grid approximation spaces for operator equations, *Math. Comp.* **78**(268) (2009), 2223–2257.
- [14] A.N. KOLMOGOROV, Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse, *Ann. Math.* **37** (1936), 107–110.
- [15] H. KÖNIG, Eigenvalue distribution of compact operators, Birkhäuser, Basel, 1986.
- [16] T. KÜHN, W. SICKEL AND T. ULLRICH, Approximation numbers of Sobolev embeddings – sharp constants and tractability, *J. Complexity* **30** (2014), 95–116.
- [17] B.S. MITYAGIN, Approximation of functions in  $L^p$  and  $C$  on the torus, *Math. Notes* **58** (1962), 397–414.
- [18] N.S. NIKOL'SKAYA, Approximation of differentiable functions of several variables by Fourier sums in the  $L_p$  metric, *Sibirsk. Mat. Zh.* **15** (1974), 395–412.
- [19] E. NOVAK AND H. WOŹNIAKOWSKI, Tractability of multivariate problems. Volume I: Linear information. EMS, Zürich, 2008.
- [20] E. NOVAK AND H. WOŹNIAKOWSKI, Optimal order of convergence and (in)tractability of multivariate approximation of smooth functions. *Constr. Approx.* **30** (2009), 457–473.
- [21] E. NOVAK AND H. WOŹNIAKOWSKI, Tractability of multivariate problems. Volume II: Standard information for functionals. EMS, Zürich, 2010.
- [22] E. NOVAK AND H. WOŹNIAKOWSKI, Tractability of multivariate problems. Volume III: Standard information for operators. EMS, Zürich, 2012.
- [23] A. PIETSCH, Operator ideals. VEB Deutscher Verlag der Wissenschaften, Berlin, 1978, and North-Holland, Amsterdam - New York, 1980.
- [24] A. PIETSCH, History of Banach spaces and linear operators. Birkhäuser, Basel, 2007.

- [25] A. PINKUS, *n*-widths in approximation theory, Springer, Berlin, 1985.
- [26] CH. SCHWAB, E. SÜLI AND R.A. TODOR, Sparse finite element approximation of high-dimensional transport-dominated diffusion problems, *ESAIM: Mathematical Modelling and Numerical Analysis* **42** (05) (2008), 777–819.
- [27] W. SICKEL AND T. ULLRICH, Smolyak’s Algorithm, Sampling on Sparse Grids and Function Spaces of Dominating Mixed Smoothness, *Jenaer Schriften zur Math. und Inf.* **14/06** (2006), 57 pp.
- [28] W. SICKEL AND T. ULLRICH, Smolyak’s Algorithm, Sampling on Sparse Grids and Function Spaces of Dominating Mixed Smoothness, *East Journal on Approximation* **13** (2007), 387–425.
- [29] W. SICKEL AND T. ULLRICH, Tensor products of Sobolev-Besov spaces and applications to approximation from the hyperbolic cross, *J. Approx. Theory* **161** (2009), 748–786.
- [30] W. SICKEL AND T. ULLRICH, Spline interpolation on sparse grids, *Applicable Analysis* **90** (2011), 337–383.
- [31] S.A. SMOLYAK, Quadrature and interpolation formulas for tensor products of certain classes of functions, *Dokl. Akad. Nauk* **148** (1963), 1042–1045.
- [32] V.N. TEMLYAKOV, About the approximation of periodic functions of several variables with bounded mixed derivatives, *Dokl. Akad. Nauk* **253** (1980), 544–548.
- [33] V.N. TEMLYAKOV, *Approximation of periodic functions*, Nova Science, New York, 1993.
- [34] H. YSERENTANT, *Regularity and Approximability of Electronic Wave Functions*, Lecture Notes in Mathematics, Vol. 2000, Springer-Verlag, Berlin, 2010.